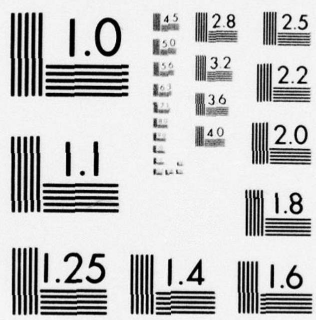
AD-A059 634 UNIVERSITY COLL LONDON (ENGLAND) DEPT OF STATISTICS --ETC F/G 17/2
THE INDRA NETWORK MEASUREMENT PROJECT.(U)
JUN 77 A V STOKES N00014-77-G-0005
UNCLASSIFIED TR-30 NL

1 OF
AD
A059634

END
DATE
FILMED
12-78
DDC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

UNCLASSIFIED ... (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: (14) TR-30
2. GOVT ACCESSION NO.
3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle): (6) The INDRA Network Measurement Project,
5. TYPE OF REPORT & PERIOD COVERED: Technical 1974-1977
6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s): Dr. A.V. Stokes (10) Adrian V. Stokes
8. CONTRACT OR GRANT NUMBER(s): (15) N-00014-77-G-~~005~~ 0005
9. PERFORMING ORGANIZATION NAME AND ADDRESS: University College London, Gower Street, London WC1, ENGLAND
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS: Defence Advanced Research Projects Agency, 1400 Wilson Boulevard, Arlington, Virginia 22209, USA
12. REPORT DATE: (11) 1 June, 1977
13. NUMBER OF PAGES
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office): Office of Naval Research (12) 93p.
15. SECURITY CLASS. (of this report): UNCLASSIFIED
15a. DECLASSIFICATION/DOWNGRADING SCHEDULE: U
16. DISTRIBUTION STATEMENT (of this Report): No.1

This document has been approved for public release and sale; its distribution is unlimited.

(9) Technical rept.,

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)
18. SUPPLEMENTARY NOTES
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Measurements of Network Usage, Packet Switched Networks, ARPANET
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report is a definitive guide to the user-level of the Network measurement project at the University College London node of ARPANET. It describes the aims of the project, describes the concepts behind the various programs and their implementation and gives the results obtained so far.

409 774

TABLE OF CONTENTS PAGE

PART I: INTRODUCTION

PART II: METHODS

PART III: RESULTS

LIST OF FIGURES

PAGE

PART I

INTRCDUCTION

1. INTRODUCTION AND SYSTEM OVERVIEW

The ARPA computer network (Ref. 15) has been operational for over six years. During this time, there have been extensive measurements on the performance of the communications subnetwork, particularly by the Network Measurement Centre at the University of California (UCLA, e.g. Refs. 8 and 10) and Bolt, Beranek and Newman (BBN; e.g. Ref. 2). There have been extensive measurements of usage of specific hosts, for example by the National Bureau of Standards (NBS; Ref.26) There have also been certain measurements of the network usage made for certain large applications to justify the cost of running the network. There have not, however, been any consistent measurements of network usage via one site. There are several reasons for this omission. Partly it is due to there being no mechanism by which US users of ARPANET could be forced to keep statistics of their usage, and partly it is due to there being no automatic accounting system for the use of the network.

An attempt was made in late 1974 by BBN to introduce an automatic accounting system into the subnetwork. The mechanism was that each communication computer would connect to a specific Access Control Host before it permitted a connection to be opened to any other Host. Further connections were permitted only if the correct user/password combination was given; after each session over a virtual circuit, the accounting Host was informed of the length of the connection and the number of packets transferred (unless the Access Control computer was not available, in which case the statistics were stored for later transmission). This mechanism was abandoned after a few weeks for several reasons; amongst these were the difficulty of maintaining the password file on the Access Control Host and the sluggishness of response of the access control mechanism.

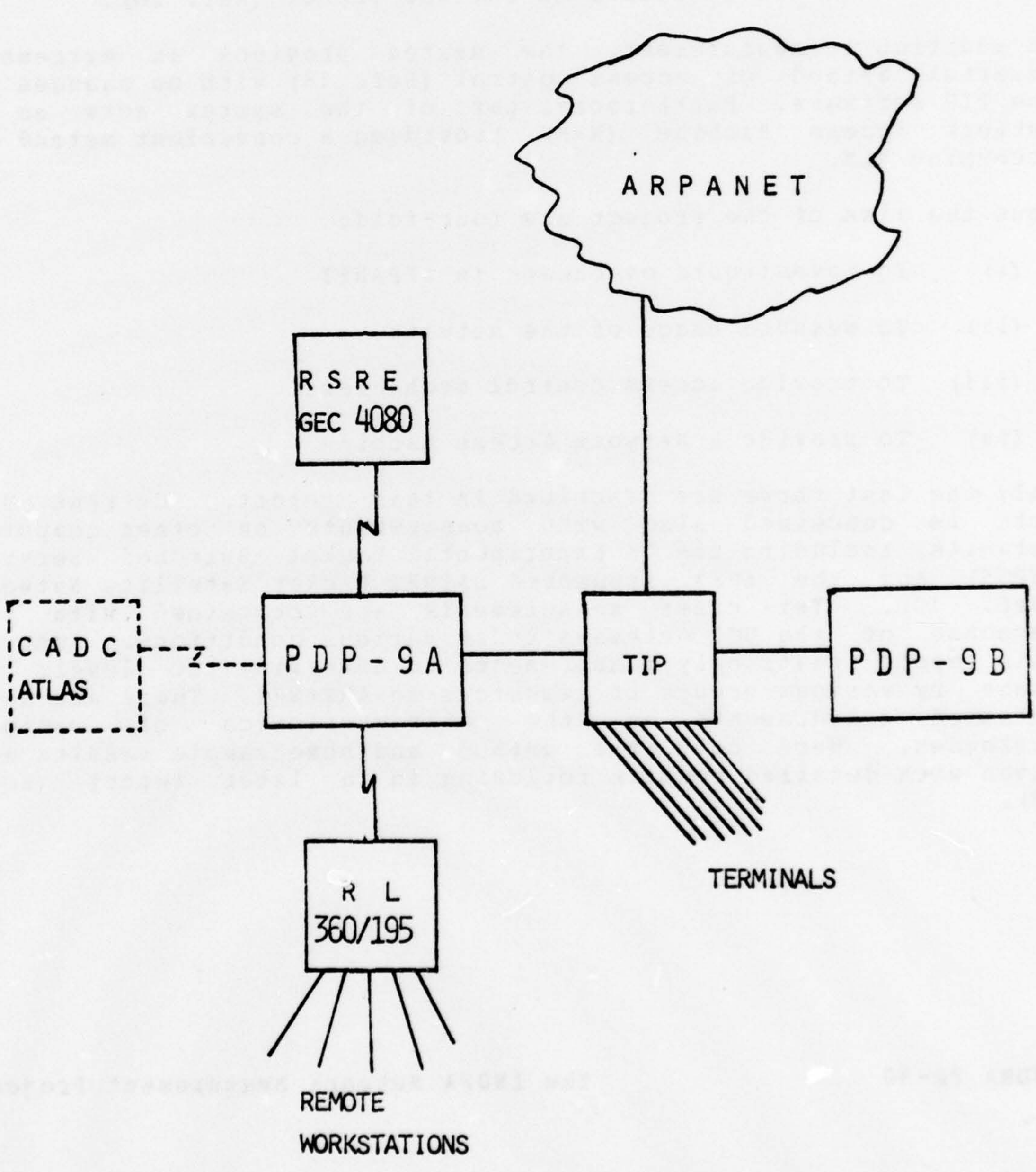
At about the same time we, at the University College London (UCL) node of ARPANET, became interested in providing access control and accounting. There were two reasons for this. Many bodies wished to analyze the extent and value of usage of the ARPANET link via the UCL node; they also wished to be able to control the access - particularly via the Public Switched Telephone network (PSTN). Because we were a node like any other, it was not possible to put any special code into the Honeywell 316 Terminal Interface Message Processor (TIP, Ref. 13) which acts as the communication processor to the UCL site. Instead, any such code had to be provided outside the subnetwork. However, since all the use of the UCL node is purely experimental, we were entitled to enforce any extra login procedures we wished onto our users. In addition to providing such access control mechanisms, we also

provided a number of measurement facilities and it is with both the access control and measurements with which this report is concerned; we also describe the network access function we have provided. Because the same system provides both measurement and access control, it is the use being made of the UCL node of ARPANET over switched telephone lines or through one of the two PDP-9s described in the next paragraph.

Two PDP-9 and one PDP-11/35 computers are connected to a TIP sited at UCL. One (PDP-9A) acts as a gateway between various computer networks, in particular, between ARPANET and the Rutherford Laboratory (RL) star network based on an IBM 360/195 (Ref. 16). On this, we are able to monitor access to these machines. The second (PDP-9B) is used both as a development machine and for monitoring and access control of the dial-up lines of the London-TIP. It also provides a simple form of access to the National Library of Medicine (NLM) Medline system. The PDP-11 is used as a gateway to a broadcast satellite network. Our current configuration is shown in Fig. 1.

This report is intended to be a definitive guide to a major part of the measurement and access control project at UCL (see Sec. 2). Part II describes the methods used in detail and presents sample results and examples of the data. Part III is a comprehensive summary of the results which have so far been obtained (over a period of about a year) and Part IV is a user's guide to the various programs developed for the project.

FIGURE 1: CURRENT INDRA CONFIGURATION



2. AIMS OF THE PROJECT

Many research groups in the UK access various Hosts of ARPANET through the London-TIP for a variety of purposes. The aim of this project is to measure that usage. This measurement is carried out at a number of levels, from the user level down to the line level. In many cases, the user level measurement consists solely of measuring the time connected to a particular Host, but in the case of access to one machine, the National Library of Medicine (NLM) IBM 370/158, much more detailed measurement is performed. At the line level, other measurements are made which are described in another report (Ref. 26).

In addition to measurement, the system provides an extremely powerful method of access control (Ref. 18) with no changes in the TIP software. Furthermore, part of the system acts as a Network Access Machine (NAM) providing a convenient method of accessing NLM.

Thus the aims of the project are four-fold:

(i) To investigate overheads in ARPANET

(ii) To measure usage of the network

(iii) To provide access control mechanisms

(iv) To provide a Network Access Machine

Only the last three are described in this report. Current UCL work is concerned also with measurements on other computer networks, including the UK Experimental Packet Switched Service (EPSS) and the ARPA sponsored SATNET Packet Satellite Network (Ref. 10). Yet other measurements are concerned with the response of the UCL Gateways under various conditions. Most of this report treats only measurements to determine the levels of usage by various groups of resources on ARPANET. There are more detailed measurements on the characteristics of various dialogues. Here only the method and some sample results are given with detailed results following in a later report (Ref. 23).

3. OTHER ARPANET MEASUREMENT

Measurement has been carried out at other ARPANET sites and it was not the purpose of this project to duplicate such work but rather to complement it. Two sites deeply concerned with monitoring are the National Bureau of Standards (NBS) and the University of California at Los Angeles (UCLA) Network Measurement Center (NMC). In this section, we mention the work which has been carried out at those sites, to point out the differences from our approach.

3.1 National Bureau of Standards

There are two projects at NBS which are relevant to our work. The first is that on the Network Measurement System, the second on the Network Access Machine (NAM). The Network Measurement System consists of two parts. The first is the Network Measurement Machine (NMM) which is (Ref. 27) "a device used to acquire data for the performance measurement of computer network systems and services". It consists of a DEC PDP-11/20 with some special purpose hardware and software. This machine performs the actual data capture and stores the data on magnetic tape for later processing by the Data Analysis Programs (DAP).

The data collection is performed in such a way as to copy the data with no interference (*e.g. a "T-connection"*). The data is recorded on magnetic tape with each character being timestamped; in addition, the source of the character is recorded. Since these data are recorded in time order, the many conversations being monitored by the NMM are multiplexed together and hence the pre-processing consists of a routine (RAW) which performs the necessary demultiplexing. The resultant conversations are then (if occurring in full-duplex mode) transformed to the equivalent half-duplex conversation. Non-printing characters are transformed into a printable form (e.g. carriage return is printed as <CR>). The resultant files are then analyzed in detail giving histograms, frequency distributions etc. In particular, the OMNITAB package (Ref. 4) is used to perform many statistical functions.

Besides the obvious relevance of the NMM to the work at UCL (and, as will be shown below, the similarity of technique), the work is of interest since one of the systems measured by the NMM is the MEDLINE system at NLM.

3.2 University of California at Los Angeles

Perhaps one of the most important functions performed on the ARPANET is that of the Network Measurement Centre. This has performed measurements on the ARPA sub-network since its inception. These measurements are specifically at the sub-network level and therefore not directly related to the results presented in this Report, although they are of relevance to the sub-network measurements of Treadwell (Ref. 26).

In addition to performing simple measurements, the NMC also has facilities for performing specific experiments (in conjunction with the Network Control Center) and these are carried out at frequent intervals. The results of such experiments enable incipient faults to be detected, allow evaluation of the various low-level protocols, give breakdowns of network traffic in terms of the various overheads etc.

An example of the results obtained is given in Ref. 12.

PART II

METHODS

1. THE USER LEVEL ACCESS AND MEASUREMENT SYSTEM

1.1 QUES Access Control

The program which performs the actual monitoring and access control is called QUES. This is one segment of a network system which runs on PDP-9B. When the system is initialized and has established communication with the TIP, QUES sets up a control connection to a specific port on the TIP; via this connection, QUES attempts to connect to other TIP ports (which are specified by entries in a disk file which may easily be modified). Each port may be in one of three states. The first is WILD, in which case there is no user connected; in this state, QUES may (and indeed does) make the connection. When a user dials in, this connection is broken and, on noting this, QUES remakes it and interrogates the user by asking for surname, TIP password and number of required Host (a typical scenario is shown in Fig. 2). To save time, a user may give all three replies in response to the first question.

If the replies are satisfactory, QUES then breaks the connection and waits 20 seconds before attempting to remake it. If the replies are unsatisfactory, the user is allowed a second attempt then, if still incorrect, he is disconnected. It would have been possible for us to have better access control by making the connection for the user. This would require, however, a considerably heavier cpu load and to reduce this, we only make the connection in the specific case that we wish to record the whole dialogue for subsequent analysis (as we do in the case of MEDLINE; see below).

If the user gives correct answers and succeeds in connecting to the Host within the requisite time, QUES enters its third phase in which it attempts (at one second intervals) to reconnect to the port.

When the user closes his connection to the remote Host, QUES is once more able to connect to him and requests the number of the next Host required (name and password are not requested again). This procedure is then repeated continuously.

```
LONDON-TIP MONITORING SERVICE.

SURNAME >stokes
TIP PASSWORD >xxx
PASSWORD UNKNOWN - REENTER >ucl
HOST NUMBER >42
OK - BYE
Closed                              * message from the TIP

@L 42                               * user logs in to host
.......
@C                                  * close connection to host

Open                                * message from TIP
NEXT HOST NUMBER OR RING OFF NOW >
```

Figure 2: A Typical QUES Scenario

All the data supplied by the user are printed by QUES onto a paper tape (thus obviating problems such as closing files after a system crash); similarly, when the user disconnects from his Host, his connect time is recorded on the tape.

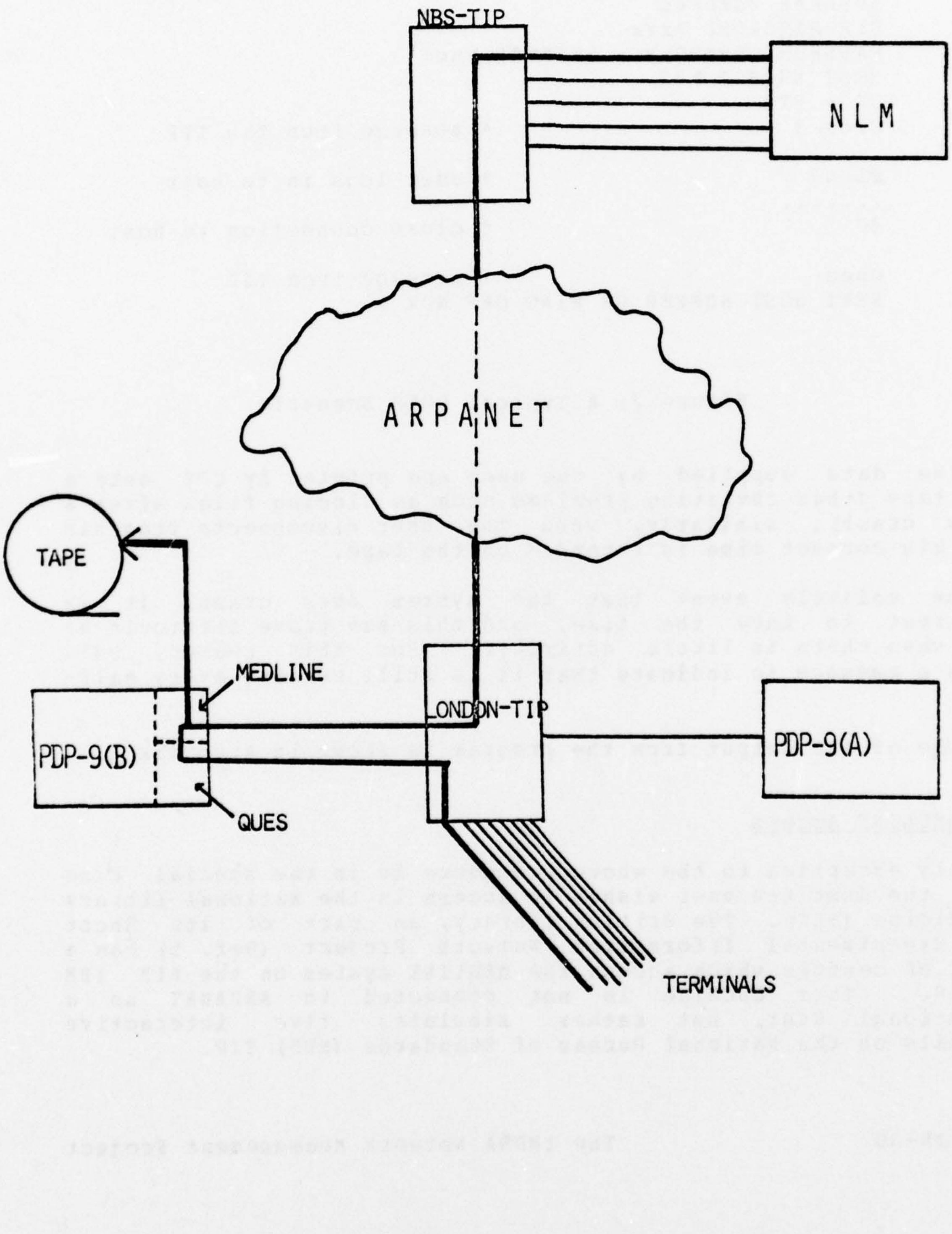
In the unlikely event that the system does crash, it is convenient to know the time, and this may prove difficult at night when there is little activity. For this reason, QUES prints a message to indicate that it is still running every half-hour.

A sample of the output from the program is shown in Appendix 1.

1.2 MEDLINE Access

The only exception to the above procedure is in the special case where the Host the user wishes to access is the National Library of Medicine (NLM). The British Library, as part of its Short Term Experimental Information Network Project (Ref. 5) has a number of centres which access the MEDLINE system on the NLM IBM 370/158. This machine is not connected to ARPANET as a conventional Host, but rather simulates five interactive terminals on the National Bureau of Standards (NBS) TIP.

FIGURE 3: DATA PATH FOR MEDLINE ACCESS



Since this makes the process of connection extremely inconvenient and also provides little status information, we have written a program which also runs under our network software on the PDP-9 and which automates the connection procedure. If a user specifies that he wishes to access the MEDLINE system (by giving the NBS-TIP number or MEDLINE when asked for Host number), he is automatically routed to this program. The data path for this mode of access is shown in Fig. 3.

The program attempts to make a connection to each of the NLM ports in turn; there are two ports (134, 136) at 10 characters a second and three (140,142 and 144) at 30 cps. If the user specifies that he has a fast terminal, the scan starts at port 134. An attempt is made to connect to each port. This attempt may be successful, in which case the user is connected to the port; it may be unsuccessful, in which case the next port is tried; or it may be partially successful. For example, the connection to NLM may be opened but the messages from NLM may not be those expected (if, say, a user had disconnected in the middle of a search) and the program attempts to correct this state, then connect the user.

Once the message

PLEASE ENTER USERID

is received, the user is asked for his NLM ID. This is usually one in the range NLL02 to NLL06 and the user does not care which is used. In this case, he types a carriage return and the program chooses one which it believes is not in use. If it is (signified by the appropriate message from NLM), the program notes this (for subsequent users) and chooses another until the user is connected. The program then prints a message to the user indicating the port and userid (in case the system crashes and the user has to reconnect manually). From this stage on, the program maintains a virtually transparent connection to NLM; some NLM system messages are modified for brevity and convenience (e.g. the PROG: prompt is removed entirely). In addition, the program is able to monitor the entire conversation if required (see the next Section).

If the UCL program is unable to make the connection (or if the user states that he does not wish to use the program by specifying the Host as NLM), QUES allows the user two minutes rather than the standard twenty seconds to make the connection since the procedure is considerably more complicated in this case.

1.3 MEDLINE Measurement

The program described in the previous section has an option which enables it to monitor all interactions. The use of the UCL node is recognised as purely experimental. All users of the UCL TIP via the switched telephone network have to agree to the conditions of such usage; these conditions include that their interactions may be recorded and analyzed. This restriction does not apply to usage of the UCL node from the US; in any case, such data cannot be captured by the UCL measurement system, except to the extent described in Section 2. The monitoring consists of copying all the input and output to or from the user and NLM, together with the output to the system console and log, onto an industry compatible 9 track, 800 bpi magnetic tape (or, optionally, onto the PDP-9 disk - this option is convenient for test runs). Each record is datestamped (to the nearest tenth of a second) and has the channel designation, together with the direction of the data, added. Thus a sample record might be:

03 MAY 76 09 52 10.1 K< TERMINAL (FAST OR SLOW)>

in which, at the specified date and time, the message "TERMINAL (FAST OR SLOW) >" was sent to channel K (which happens to be a user on TIP port 70). The system console and the system log are represented by the characters # and * respectively. A sample of the data recorded on tape is shown in Appendix I.

Two other items are noticeable in this data. The first is that the data from all the channels is multiplexed together and hence must be demultiplexed before it can be analyzed (see Sec. 3). Secondly, some messages are sent with no terminating carriage return/line feed and hence these characters are appended to each such line, together with a special character ($) to denote that this has been done.

The data is recorded on the tape with one file corresponding to one session (where a session is terminated either by the operator or by a system crash). The tape is then sent to the RL 360/195 for analysis as described in Sec. 3.

2. GATEWAY MONITORING

The previous section has described the use made of PDP-9B. This section describes measurements recorded on PDP-9A. This machine is attached as a local host to the TIP and acts as a gateway between ARPANET and various other computers. In particular, at the present time, it is connected to the RL 360/195 (which is the central node of a star network, the other nodes being, in the main, GEC 2050 computers simulating IBM 1130s) and the Royal Signals and Radar Establishment (RSRE) GEC 4080. Previously, it was connected to the Cambridge Computer Aided Design Centre (CADC) Atlas computer.

The RL 360/195 is one of the largest computers on ARPANET and hence its use from the network is of considerable interest. This use has been restricted to specific collaboration between research groups in the US and RI, together with local use by UCL staff who access the 360 through the TIP. There has been a small amount of other usage. Thus the usage is, in a sense, the inverse to that measured by QUES described above, although usage directly corresponding to QUES is provided by RL users accessing ARPANET through the PDP-9. The RSRE 4080 is used by research groups in the US accessing the CORAL compiler, mainly for evaluation purposes.

The data recorded by PDP-9A is also on a paper tape for the reasons outlined in the previous section. This tape is a direct transcript of the system log and records such information as the remote computers (and, of course, the IMP) being up or down, the use by various people (the Host from which access was made, the identifier used and the initials of the person) and, since these data are timestamped (only to the nearest minute), the time connected to the various machines. In addition, any batch jobs (including those arising from file transfers) are also recorded. Sample data are shown in Appendix 1.

3. ANALYSIS PROGRAMS

The analysis programs for the user level analyses on the RL 360/195 are combined into one module called XFSTATS. All the programs are written in Babbage (Refs. 15 and 25) and make heavy use of its system library (Ref. 19).

The module consists of a central driver routine, together with a number of independent segments wich perform various analysis functions. The largest one, QUES, analyzes the data produced by the PDP-9 QUES program and produces various matrices, histograms etc.

3.1 Overview of the XFSTATS package

XFSTATS takes three inputs. The first is the tables; these specify the fault captions, the allowed control cards, the legal TIP passwords, the names of the computers/networks connected to PDP-9A and the list of 360 IDs used by separate projects. The second is the data; this is effectively a concatenation of all possible inputs. The last is the control file; this consists of a sequence of lines, each specifying an option. The order of these lines is not significant since they are not interpreted in order of reading but rather the entire file is read by a parsing program (Ref. 19) which constructs an appropriate data structure which is handed to XFSTATS. An example of a control file is:

```
QUES
FROM 1/6/75
TO   1 July 75
CONNECT TIMES
HISTOGRAM OF PORT USAGE
TIMES OF MONITORING
END
```

It can be seen that a very flexible input format is allowed. Each command may be abbreviated to four characters if required.

XFSTATS determines the type of analysis to be performed from the control cards and, where relevant, limits such as start and end dates (which default to 1 July 75 and 1 July 76 respectively). It then reads the data file, ignoring any data not within the specified limits and performs the appropriate analyses.

3.2 Facilities Provided

The XPSTATS program allows many analysis functions to be performed. In this section, we describe the facilities available. A similar description, but written from the point of view of the user, is given in Ref. 20.

(a) QUES

This is the major PDP-9B analysis program. It produces measurement data for the output from QUES. It operates by building a data structure in core for the compressed data (see below) and then, depending on the control cards, prints certain data. The functions available are:

- Times monitored
- Surnames used for each ident
- Number of logins, connect time and number of hosts used for each ident
- Number of logins, connect time and number of users for each host
- Matrix of connect times for host number/ident pairs

Where appropriate, the data may be requested in graphical form. Examples of the output are given in Appendix 1.

(b) LOGB

This program checks the input (the PDP-9 paper tape log) for errors, removes superfluous lines (e.g. "QUES OK"), removes lines where the user has given an incorrect password or Host (these are printed out for inspection, but not passed onto the analysis programs) and compresses the data. In particular, the connect time is appended to the message generated when the user logs in. Two problems arise. The first is that monitoring may be terminated while a user is still logged in; in this case, the QUES program prints out the connect time up to the termination and so the time recorded is an underestimate. This of course presents no problems to LOGB, since there is no difference between such a message and a genuine log out. The second case, when the system crashes while users are logged in, does present a problem; in such a case, LOGB detects this by the occurrence of a "MONITORING STARTED" message not preceded by a "MONITORING TERMINATED" message. It generates the latter and the associated logout messages at the last time for which it had a valid message (hence the reason for the production of the "QUES OK" messages).

The program also produces messages indicating the number of ports in use. Due to a temporary restriction on the number of channels available in the PDP-9, QUES only monitors six ports at present. It is a simple matter for LOGB to record the number of ports in use except for one case, when QUES starts monitoring. In this case, ports may already be in use; the version of QUES to which we are referring did not attempt to distinguish whether the port was in use by a genuine user or not (e.g. someone having dialled the TIP number by mistake). This distinction can be deduced by QUES with a reasonable degree of certainty by knowing the TIP timeout period. However, the timeout period is not guaranteed to remain constant and the method is not completely reliable. In the data we present here, this was not done and LOGB had to deduce the number of ports in use. Due to the mode of operation of the telephone system, a user is allocated to the lowest numbered free port on dialling up. At present, we monitor ports 70 to 75 (octal). If the first port to be used after QUES was initialized was 71, LOGB would assume that port 70 was in use at that time. LOGB takes steps to ensure that such an error is not propagated if the port was not in genuine use.

Thus LOGB produces an output file, an example of which is given in Appendix 1.2 (the output is that produced with the data in Appendix 1.1 as input). This file is in a standard format and may be assumed to be error free and it may be analyzed by QUES.

(c) IDENT and HOST

These two programs perform very similar functions and simply edit the data, only printing interactions which appertain either to a specific ident (TIP password) or host number. Strictly, they should not be included in this document since they perform no analysis functions, but it is often convenient to examine the data visually and these programs facilitate this. In addition, they print appropriate histograms if required; in the case of IDENT, the histograms are those of time and users, in the case of HOST, only the former.

(d) PORT USAGE STATISTICS

This program operates on the compressed data from either PDP-9. It produces statistics on port usage, in tabular form and graphically. The statistics produced are the time monitored, the number of logins, the amount of time used and, from these, the average time per login and the percentage TIP port utilization. The statistics may be produced over any specified dates and they are broken down by hour of day or weekend. In addition, a table

is produced indicating the number of ports in use whenever there is a new login. If used on the PDP-9A log, the program only examines the port utilization for one computer (PDP-9A, RL 360 or RSRE 4080) and the required one must be specified as an additional parameter.

(e) MEDPRINT

Although the major MEDLINE measurement is from the output on magnetic tape, some data is produced on the PDP-9B log and it is this which is analyzed by MEDPRINT. For each interaction, the program calculates the time per line from NLM and the user, the average number of characters per line each way and the ratio of number of lines from NLM to number of lines from the user. At the end, a summary of the results, including the mean and standard deviation of each value is printed.

(f) DEMUX

As described in Sec. 1.3, the various interactions monitored by MEDLINE are multiplexed together onto magnetic tape. It is the purpose of DEMUX to demultiplex them into the separate conversations. The output is then identical to that printed at the user's terminal with the addition of a datestamp to each complete line (N.B. if the user replied to a question on the same line as the question, his reply is not datestamped). In addition, various data about the interaction which are extracted from the log stream (such as the user's name and password and the number of lines and characters to and from the terminal) are added.

DEMUX may be used to analyze entire files (where a file, in general, corresponds to one measurement session) or to extract specific interactions from the file. This latter facility is of considerable use since MEDLINE monitors ALL interactions and the data on the channels connected to NLM are very similar (although not identical - see Sec. 1.3) to those to the user.

(g) MEDLINE

This is the program which performs the actual analysis of the user conversations with NLM. Once again, many options are available and the user of the analysis package may choose which are required.

The program takes an individual interaction and perform various

analyses on it. The available data are as follows:

Number of characters to and from NLM
Number of characters to and from the user
Occurrence of NLM system keywords
Number of searches in the session
Number of terms per search
Number of sub-terms per search statement

(h) LOGA

The output from the PDP-9A is a direct transcript of the console log and, as such, contains messages from various places such as the supervisor, as well as the measurement messages in which we are interested. In addition, the date and/or time are suppressed in some cases (only if they are identical to the last date or time printed). Since the format is not that expected by XPSTATS, this is cleaned up by a stand-alone program called TIDYA. This removes all blank lines and lines consisting solely of a system prompt character (>). It then removes all lines which do not conform to the expected format and prints them out on a log stream for inspection. The remainder then have the date and/or time added where required. The output from this is similar to that produced by PDP-9B and is processed by a program called LOGA.

This operates in a very similar manner to that of LOGB described in (a) above and generates an output of a similar, compressed, format (see App. 1.8). However, it is considerably more complex due to the larger number of different types of interaction possible (e.g. terminal access to 360, terminal access from 360, file transfers etc.) and the number of connected computers. The techniques used to perform the compression are described in detail in Ref. 22. This output is assumed to be error free and may be analyzed by the program PDP9A described in the next section.

(i) PDP9A

This is the major analysis program for the PDP-9A. It acts in a very similar way to that for PDP-9B (QUES, see (a) above). That is, it reads the compressed data file, generates a data structure then, depending on the analyses requested, prints out suitable tables and graphs.

The analyses available are as follows:

Times monitored
Up time of the connected computers
Number of ARPA Protocol errors by Host
Number and distribution of file transfers
Initials used for each ident
Number of logins and connect time for each ident
Number of logins and connect time from each host
Matrix of connect times for each host number/ident pair

If required, the various idents may be grouped together by project and the same data produced. In addition, the data may be produced either for ARPANET users accessing the RL 360 or 360 users accessing ARPANET. In all cases, histograms may be requested where appropriate.

4. SAMPLE RESULTS

4.1 QUES Monitoring

In a project such as this where a large amount of significant data is collected, it is not easy to give selected results. However, this has been done and the results are presented in Part III. This section is intended merely to show the type of output which the analysis routines are able to produce and to give examples of other information which may be deduced.

In all these sample results, the month of July 1975 is chosen. In the full results presented later, the year July 75 to June 76 is used. Before July 1975, although TIP password checking was performed, illegal access was not prohibited. Thus the validity of any results before this date is questionable.

A single month provides a reasonable volume of data covering various times of day. July 1975 was chosen specifically since it was the month in which we did most monitoring; PDP-9B is also used for system development and monitoring is only performed when it is not used for that purpose. Thus the amount of monitoring varies considerably from week to week and much of the monitoring is at night and weekends with few users; at such times, the main value of the system is for controlling improper access. In the month concerned, we monitored the TIP for 525 hours (70%) broken down in the following way:

	0000-0800	0800-1300	1300-1800	1800-2400	Weekends
Total(mins)	9120	5571	1713	4426	10674
Mins/day(av)	397	242	74	192	1334
% of time monitored	82.7	80.7	24.8	53.5	92.7

The three aspects with which we will concern ourselves are:

(i) The global picture; overall usage and general statistics

(ii) The pattern of usage by one specific user

(iii) The pattern of usage of one specific host

On the global picture, we monitored the TIP for 31,504 minutes (70%) in the month. The total time users were connected to various Hosts was 10,761 user minutes, giving a usage of the six ports monitored of 5.7% However, much of the time monitored was at weekends and night and the breakdown of this figure over the time periods used above is:

	0000-0800	0800-1300	1300-1800	1800-2400	Weekends
Time*ports	94	8039	1443	462	723
% ports used	0.2	24.1	14.0	1.7	1.1
Logins	2	435	62	42	44

(where the percentage referred to is the percentage port utilization). The high figure in the 0800-1300 time is due to the non-availability of many US hosts to us in their prime shift (1300 onwards, UK time). This restriction applies to the two Hosts that were used most heavily, NLM and the Information Sciences Institute (ISI) PDP-10X. A general matrix is produced of the usage at each Host by each user group. The complete matrix in our case would be quite unreadable. A partial matrix is shown in Appendix 1.5. The two most heavily used Hosts during this period were Host 147 (71%) and ISI (13%). The reason that the former was used so heavily was that it had been unavailable for most of June; in addition, the general university usage was low in July because of the onset of holidays.

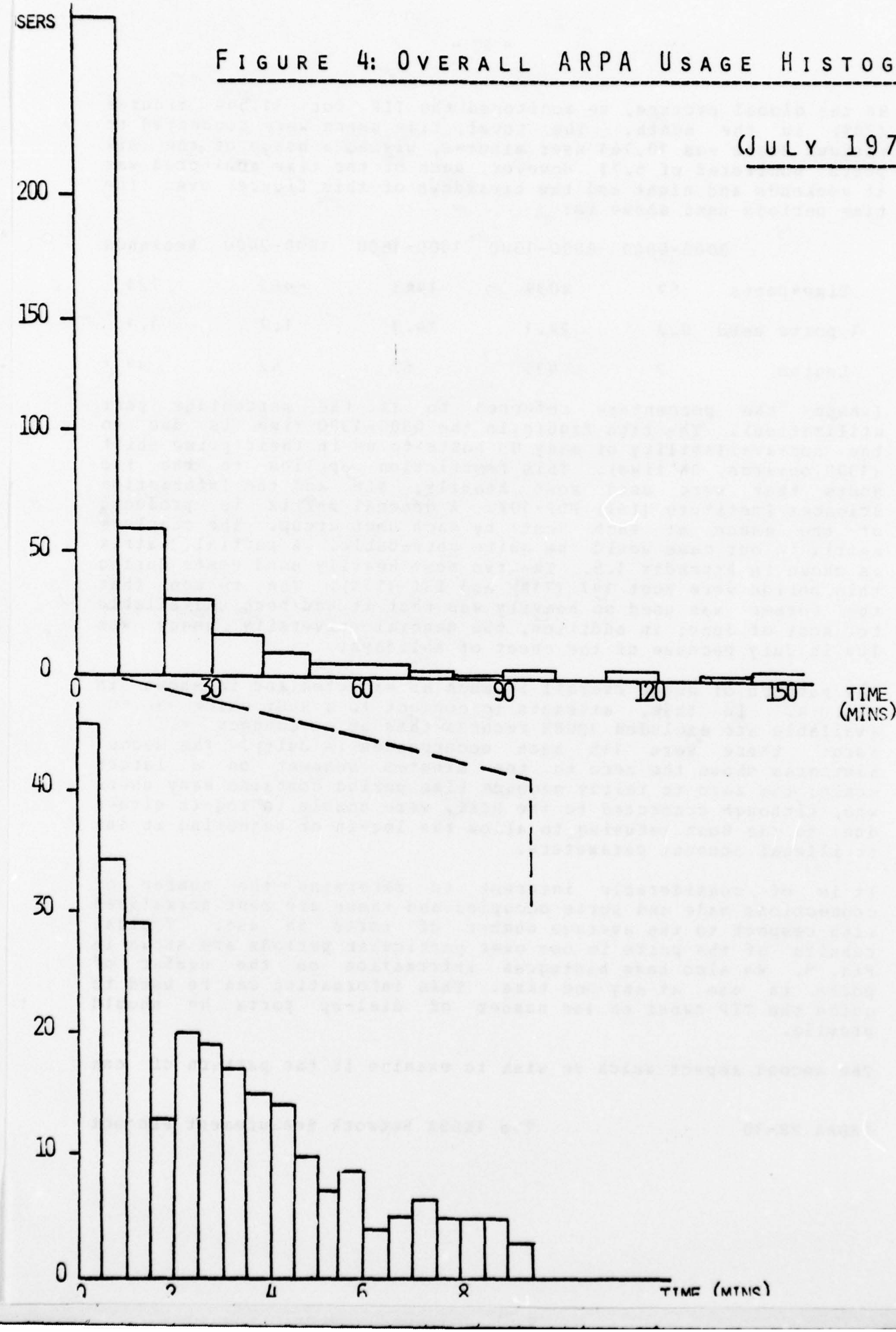
The pattern of usage overall is much as expected and is shown in Fig. 4. In this, attempts to connect to a Host which was not available are excluded (QUES records this as a connect time of zero; there were 115 such occurrences in July). The second histogram shows the zero to ten minutes segment on a larger scale; the zero to thirty seconds time period contains many users who, although connected to the Host, were unable to log-in either due to the Host refusing to allow the log-in or rejecting it due to illegal account parameters.

It is of considerable interest to determine the number of connections made and ports occupied and these are best normalized with respect to the average number of ports in use. Typical results of the ports in use over particular periods are shown in Fig. 5. We also have histogram information on the number of ports in use at any one time. This information can be used to guide the TIP owner on the number of dial-up ports he should provide.

The second aspect which we wish to examine is the pattern of one

FIGURE 4: OVERALL ARPA USAGE HISTOGRAM

(JULY 1975)



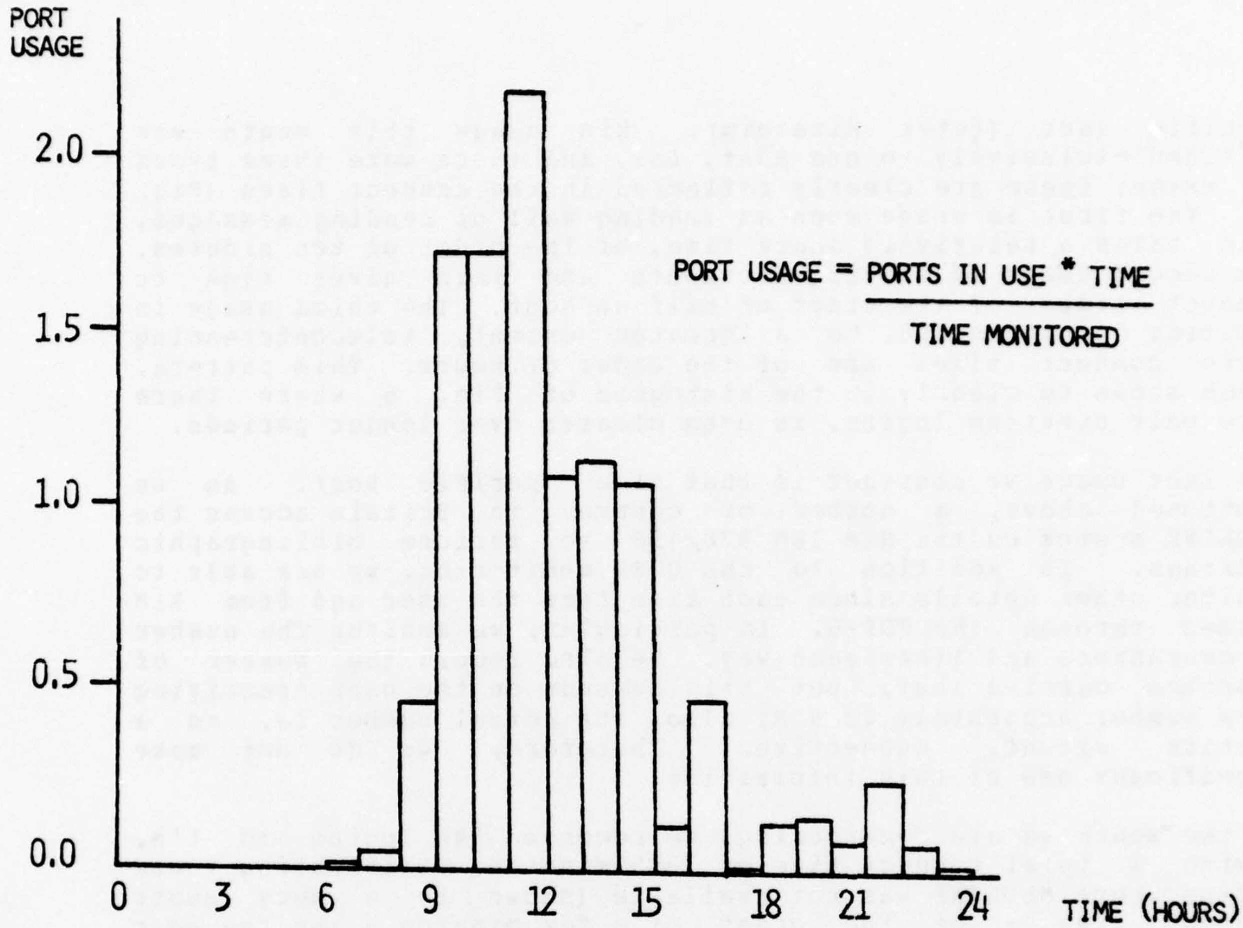


FIGURE 5: TIP PORT USAGE (JULY 1975) AS A FUNCTION OF TIME OF DAY

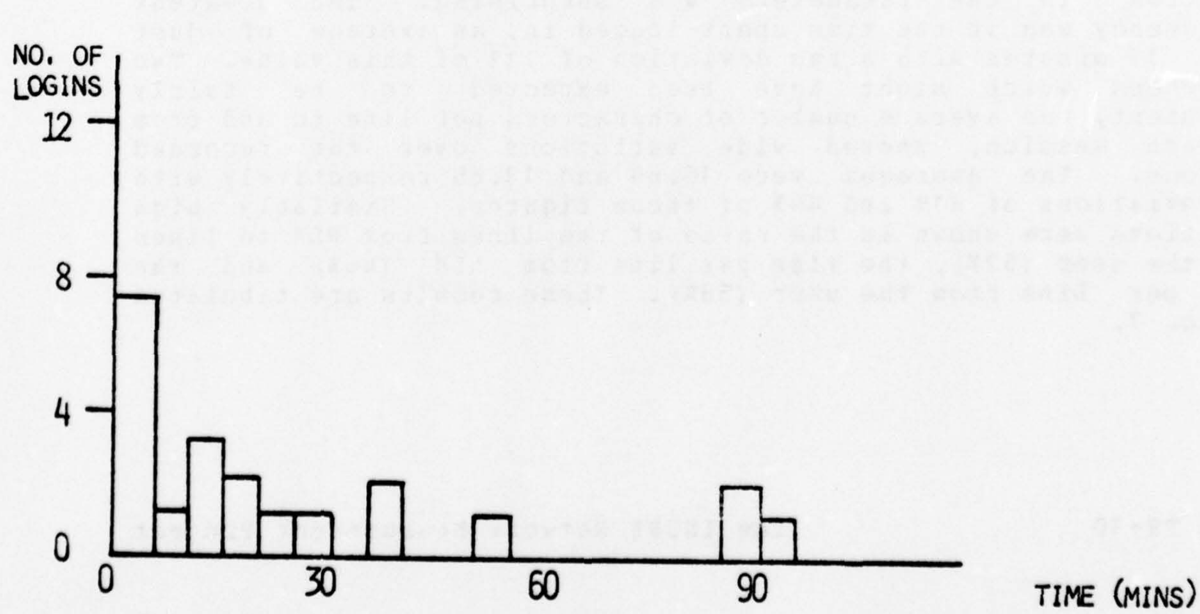


FIGURE 6: PATTERN OF USAGE FOR A SPECIFIC USER

specific user (Peter Kirstein). His usage this month was confined exclusively to one Host, ISI, and there were three types of usage; these are clearly reflected in the connect times (Fig. 6). The first is usage such as reading mail or sending messages. This takes a relatively short time, of the order of ten minutes. The second usage is editing documents and this gives rise to connect times of the order of half an hour. The third usage is entering documents and, to a greater extent, teleconferencing where connect times are of the order of hours. This pattern, which shows up clearly in the histogram of Fig. 6 where there were only nineteen logins, is even clearer over longer periods.

The last usage we consider is that of a specific Host. As we mentioned above, a number of centres in Britain access the MEDLINE system on the NLM IBM 370/158 to perform bibliographic searches. In addition to the QUES monitoring, we are able to monitor other details since each line from the user and from NLM passes through the PDP-9. In particular, we monitor the number of characters and lines each way. We also record the number of searches carried out, but this depends on the user specifying this number accurately to NLM; also, the actual number is, to a certain extent, subjective. Therefore, we do not make significant use of this information.

In the month we are considering, we recorded 148 logins to NLM, giving a total connect time of 7385 minutes. Eliminating those logins where MEDLINE was not available (shown by a very short connect time - of the order of a few minutes - and the user specifying that no searches were done and giving no "PRINT" commands), or where we did not generate full statistics, we had 87 logins.

Although the number of observations is reasonably high, the variation in the parameters was surprising. The greatest consistency was in the time spent logged in, an average of just under 33 minutes with a rms deviation of 11% of this value. Two parameters which might have been expected to be fairly consistent, the average number of characters per line to and from NLM each session, showed wide variations over the recorded sessions. The averages were 16.64 and 13.65 respectively with rms deviations of 31% and 49% of those figures. Similarly high deviations were shown in the ratio of the lines from NLM to lines from the user (52%), the time per line from NLM (46%) and the time per line from the user (58%). These results are tabulated in Fig. 7.

	Total	Average	RMS Deviation
Time logged in	171 349	1970	217
Characters from NLM	551 925	6344	5319
Lines from NLM	32 569	374	284
Characters from user	105 979	1218	2088
Lines from user	8 554	98	171
Average time per line from NLM	-	5.7	2.6
Time per line from user	-	42	24
Characters per line from NLM	-	17	5
Characters per line from user	-	14	7
Lines from NLM/Lines from user	-	8	4

Figure 7: Characteristics of NLM Usage (times in minutes)

4.2 PDP-9A Log Results

As in the previous section, we present results here to illustrate the type of data analysis possible and this may best be done with a relatively short period, namely a month. The month of February 1976 was chosen since this is the first month for which we have full monitoring since the change of output format to the PDP-9 log (the programs currently accept the new format alone, although they will be changed to allow the previous format; the change of format was simply to reduce the amount of data punched while retaining all the information).

The PDP-9 was up for over 93% of the available time. It was taken down correctly 6 times and crashed (or was taken down incorrectly) 53 times, giving a mean time between failure of over 12 hours. The IMP (as usual) was exceptionally reliable, contact being lost a total of 15 times in the month, giving a MTBF of nearly 40 hours. The RL and RSRE machines were noticeably less reliable.

The average PDP-9 port usage (3.33%) is quite low, but there were a total of 2262 logins in the month. However, only on one occasion were all five ports in use (for a total of two minutes). There is no record of the number of times a user was rejected, but obviously it cannot be more than the above figure and hence it is safe to say that this did not happen in the month.

The average usage of the three 360 ports (3.11%) was relatively low, but since usage is heavily concentrated in peak hours, this figure does not give a true indication of the usage. Perhaps a better indication is the fact that a user was rejected from using the 360 seven times in the whole month, while all three 360 ports were occupied for a total of 50 minutes (on 25 separate occasions). Therefore, in the month, the 360 usage was relatively light. The average time per login was just over 16 minutes, although this varied considerably over the time of day; with the fairly low number of logins (452 in the month), little significance can be attached to this result.

A total of 49 idents were used from 28 different ARPA hosts, although the predominant usage (61%) was, as expected, from the London-TIP. It should be emphasised that the local usage is very seriously underestimated since the HASP system console is a permanently connected device which does not go through the TIP and the usage of this device (which is high) is not recorded by the logging program.

We are able also to record the usage of ARPANET from the 360 and, in the month in question, a total connect time of 834 minutes was recorded (cf. 14675 minutes in the other direction). This usage occurred from 16 different devices attached to the 360, some directly attached, others via workstations. The most heavily used hosts were BBN and ISI (presumably for message traffic) and Lawrence Berkeley Laboratories which have a joint high energy physics experiment with RL.

Thus, to briefly summarise the results, the machines were available for a good proportion of the time, the usage of the PDP-9 and RL ports was fairly low and the usage of the 360 from ARPANET was very much more than the usage in the reverse direction.

PART III

RESULTS

1. Results of Monitoring PSTN Usage

1.1 Introduction

In this report, we will report on data obtained up to the end of June 1976. For reasons explained before (i.e. no password validation), data before July 1975 cannot be considered valid. Therefore, we have a year's statistics and it is results from these data which we present in this section. In brief, the results are those shown above in the sample results, but for a year instead of a single month. In addition, of course, variations from month to month are noted. The data we record are essentially the time monitored, the hosts used by each group and the time connected to each host. There are many ways in which such data may be analyzed and, in this report, we present those results which we consider of most interest. The full results (the actual 360 output) are given in Ref. 21. We break down the results we give here into several distinct sections.

1.2 Monitoring Performed and TIP Port Usage

The monitoring performed varies significantly from day to day depending on the load on the PDP-9 and, although the overall monitoring is of the order of 50%, much of this is at night and hence of little interest since usage is extremely low. The monitoring required a dedicated PDP-9 and it could not be spared for this purpose during much of the working day. Fig. 8 shows the amount of monitoring performed by month and Fig. 9 is the monitoring over the period broken down by hour of day (or weekend). Fig. 10 shows the frequency of port usage; it is a histogram where the value measured is the number of ports in use each time there is a new login.

The amount of port usage, broken down by month, is shown in Fig. 11. Although the figures appear very low, they are explained by the preponderance of monitoring during the night. This is shown more clearly in Fig. 12 which breaks down port usage by time of day. Nevertheless, the maximum port usage, just before lunch time, is only about 20%. Fig. 12 is mirrored closely by Fig. 13 which records the number of logins per hour - again a very low figure. In order to attempt to remove the consistent bias introduced by monitoring only in the mornings and off-peak periods, a full one week experiment was carried out in mid-February 1977. The results of this one week have not yet been analyzed but will be included in Ref. 24.

The monitoring performed allows us to give an accurate estimate

FIGURE 8: QUES MONITORING PERFORMED

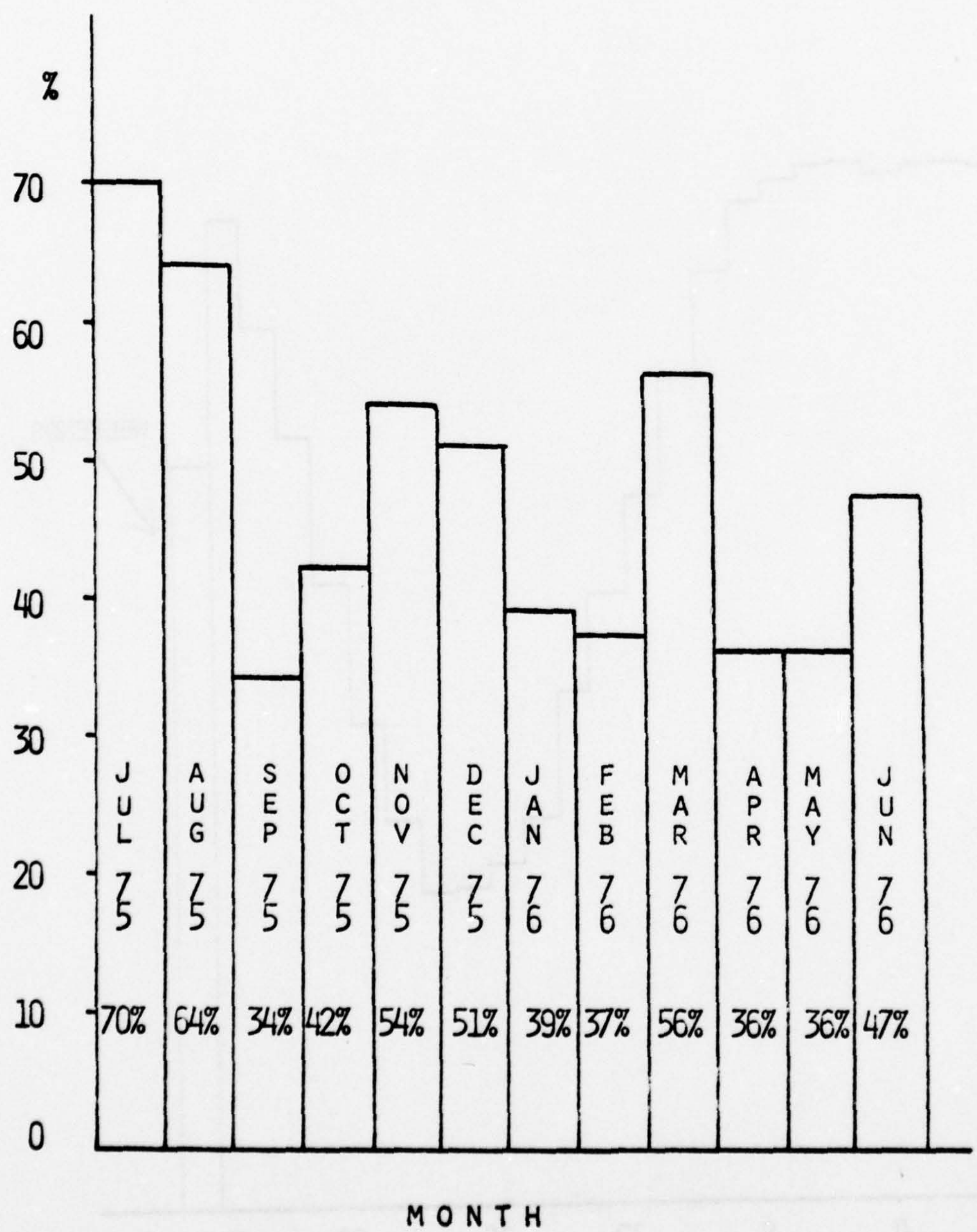


FIGURE 9: MONITORING PERFORMED AS A FUNCTION OF TIME OF DAY

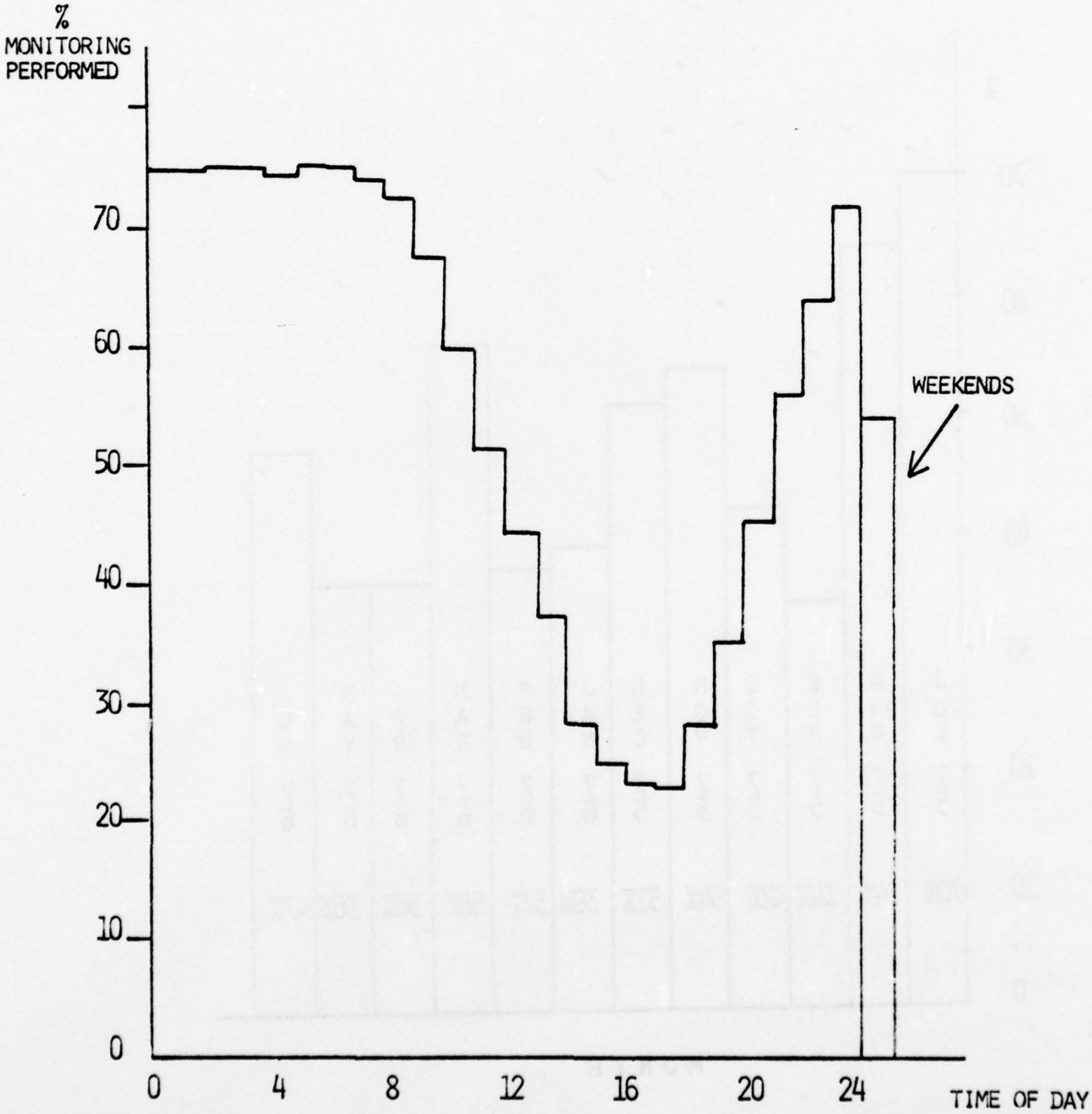


FIGURE 10: FREQUENCY OF PORT USAGE

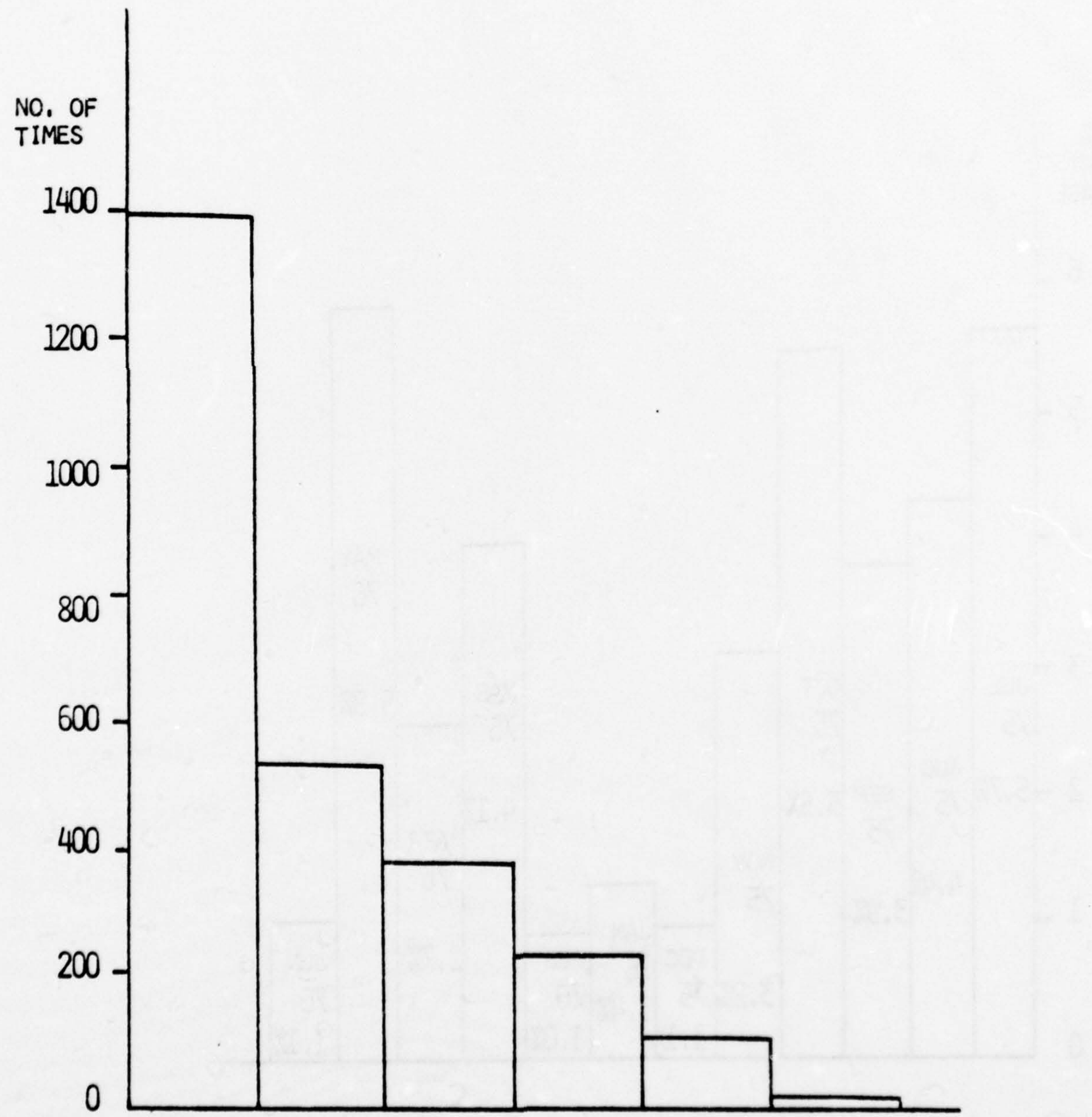


FIGURE 11: PORT USAGE BY MONTH

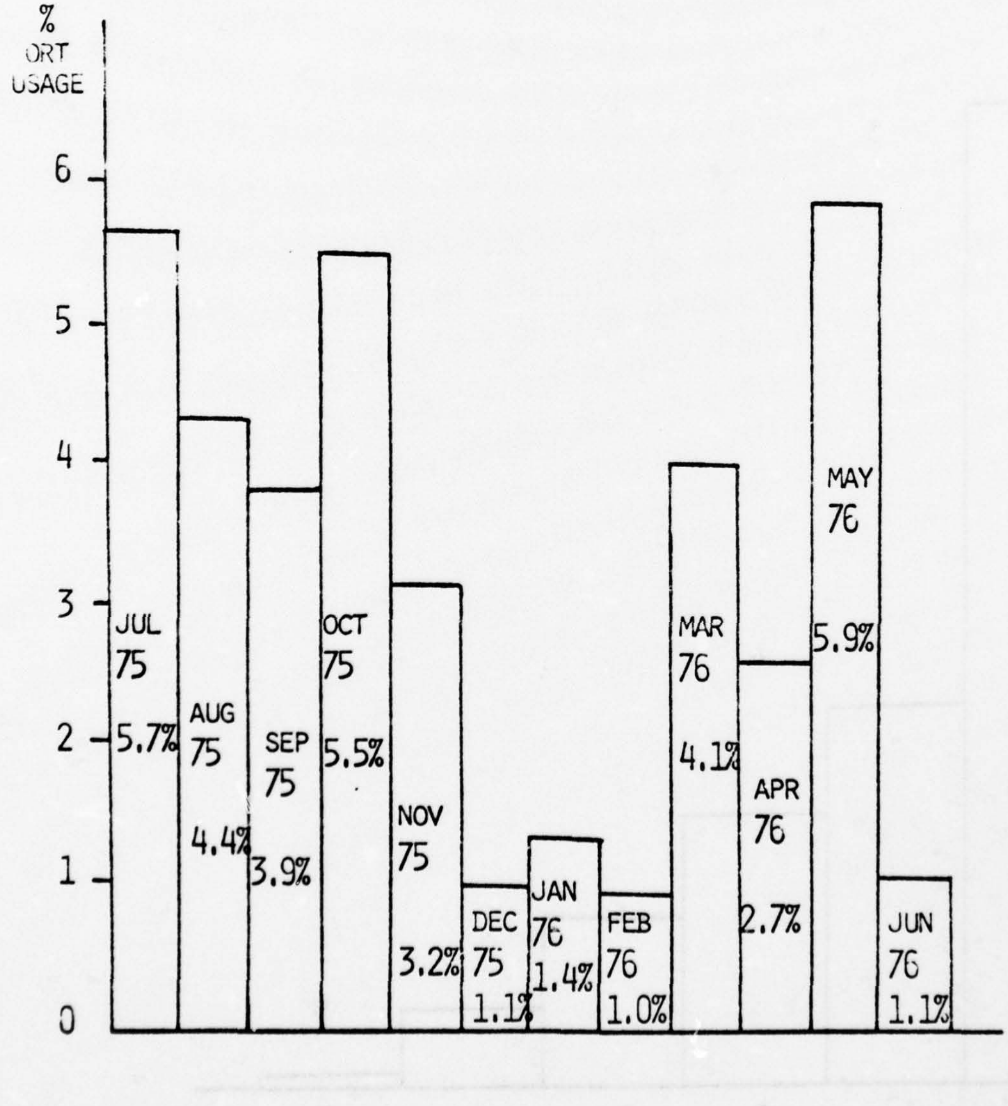


FIGURE 12: PORT USAGE BY TIME OF DAY

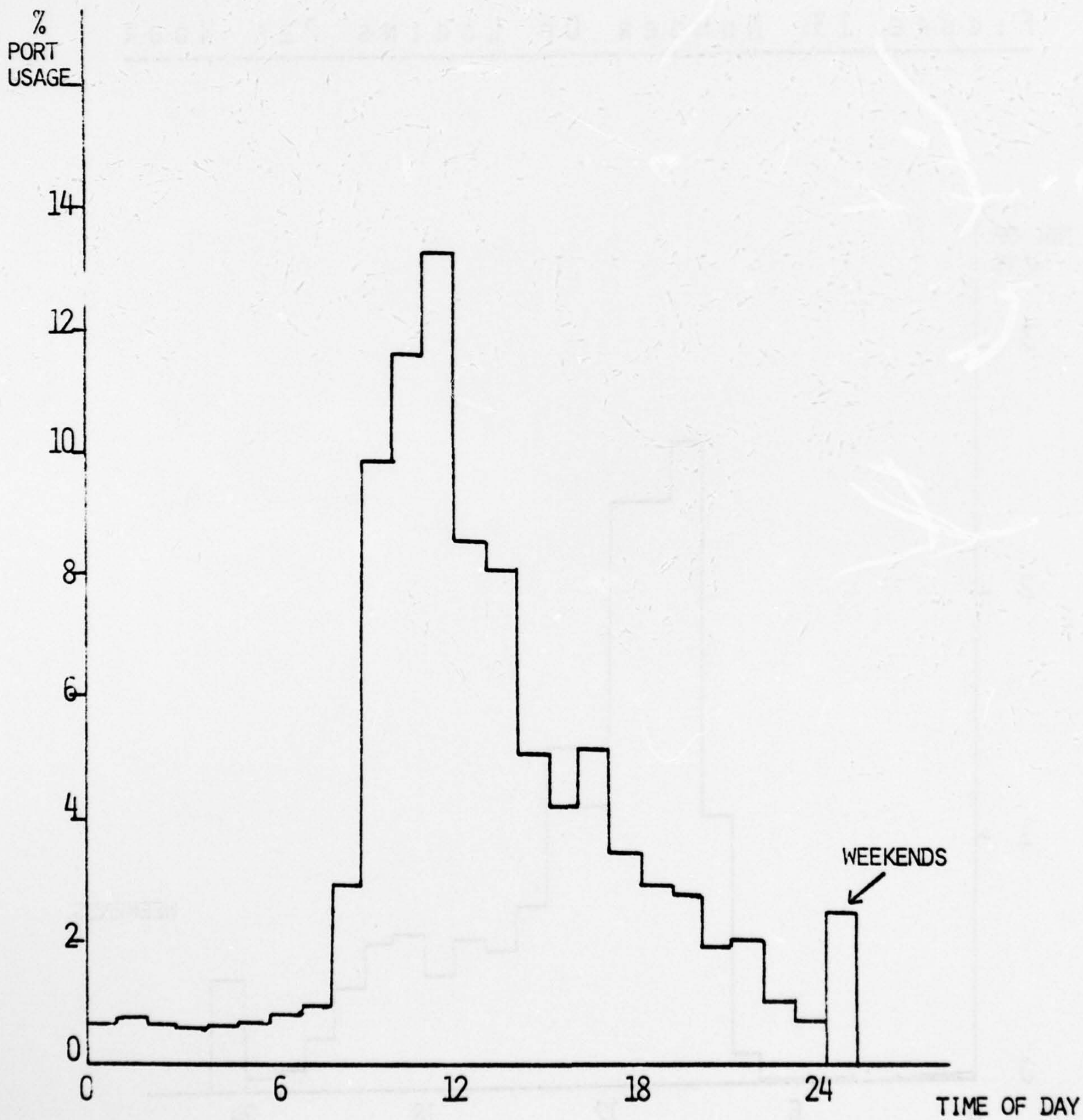


FIGURE 13: NUMBER OF LOGINS PER HOUR

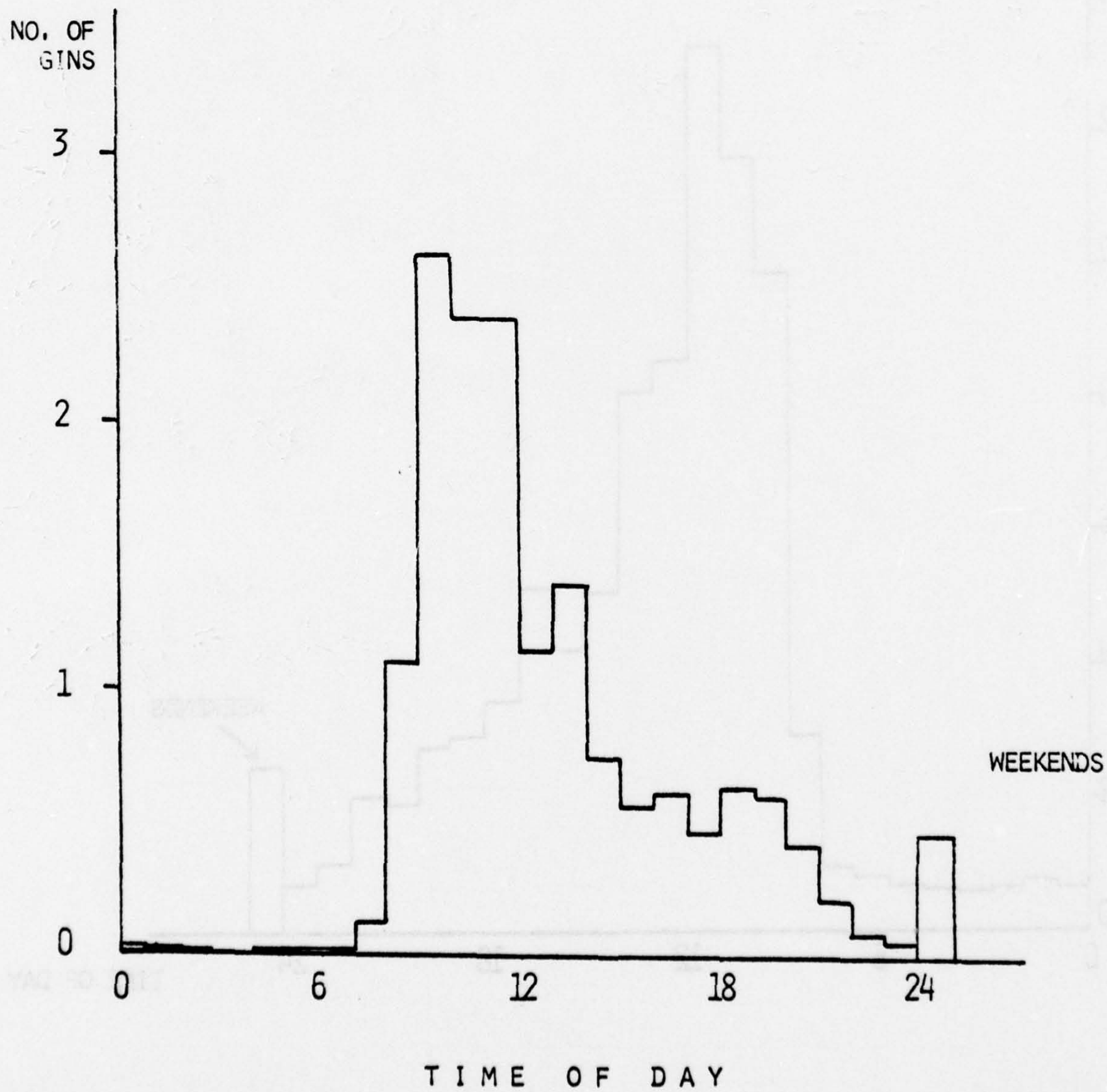
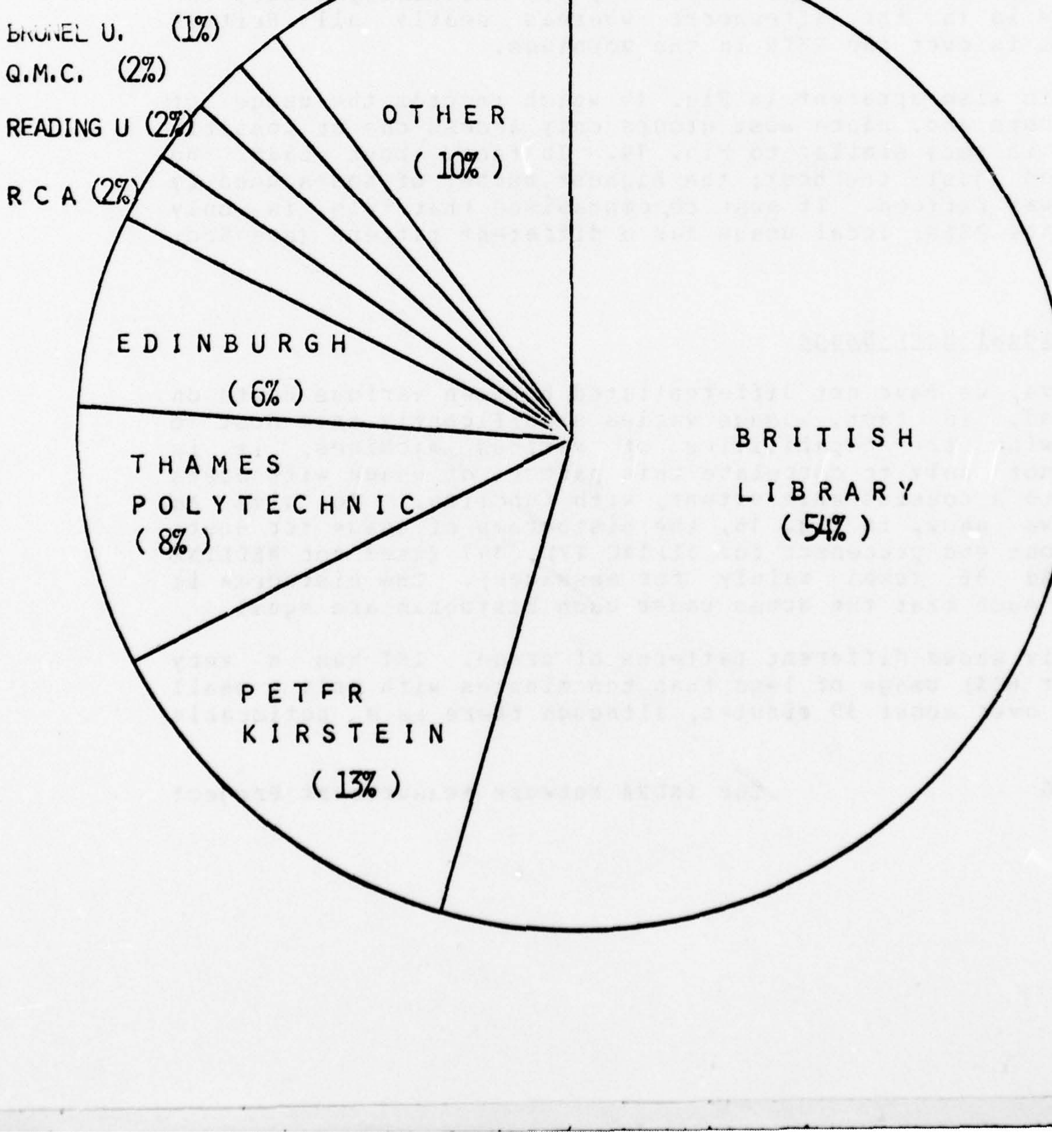


FIGURE 14: USAGE OF ARPANET BY VARIOUS RESEARCH GROUPS



of port usage; partly as a result of these measurements, we have recently cancelled some of the dial-up lines to the London-TIP. This move related partly, however, to a decision to have a substantial proportion of PSTN usage passing through EPSS.

1.3 Global ARPANET Usage

The data for usage, broken down by host and user group is available in the form of the matrix as shown in App. 1.5. The full matrix is far too large to reproduce (42 x 53) and the data may best be presented in the form of a "pie chart". The chart for the various user groups is given in Fig. 14. This shows very heavy usage by the British Library (MEDLINE and CANCERLINE Projects), over 54% of total usage. There are two reasons for this; first, there was fairly heavy usage in the period because there had been problems before the period we are discussing and this very significantly increased usage over the summer when activity from other groups was very low; secondly, QUES only records usage via the PSTN which heavily biases the data since considerable usage is made of the network via directly connected terminals or leased lines. Moreover, for university usage, the major usage is in the afternoons whereas nearly all British Library use is over the PSTN in the mornings.

This bias is also apparent in Fig. 15 which records the usage of various hosts and, since most groups only access one or possibly two hosts, is very similar to Fig. 14. In fact, about half the groups used just one host; the highest number of hosts used by one group was fifteen. It must be emphasised that this is only usage via the PSTN; local usage has a different pattern (see Sec. 3).

1.4 Individual Host Usage

In the above, we have not differentiated between various Hosts on ARPANET and, in fact, usage varies significantly from Host to Host. Knowing the capabilities of various machines, it is possible not only to correlate this pattern of usage with Hosts but also, to a considerable extent, with function. To give an example, we show, in Fig. 16, the histograms of usage for Hosts 16 (the front end processor for ILLIAC IV), 147 (used for MEDLINE access) and 86 (used mainly for messages). The histogram is normalized such that the areas under each histogram are equal.

This clearly shows different patterns of usage. ISI has a very high (over 60%) usage of less than ten minutes with only a small percentage over about 30 minutes, although there is a noticeable

FIGURE 15: USAGE OF VARIOUS ARPANET HOSTS

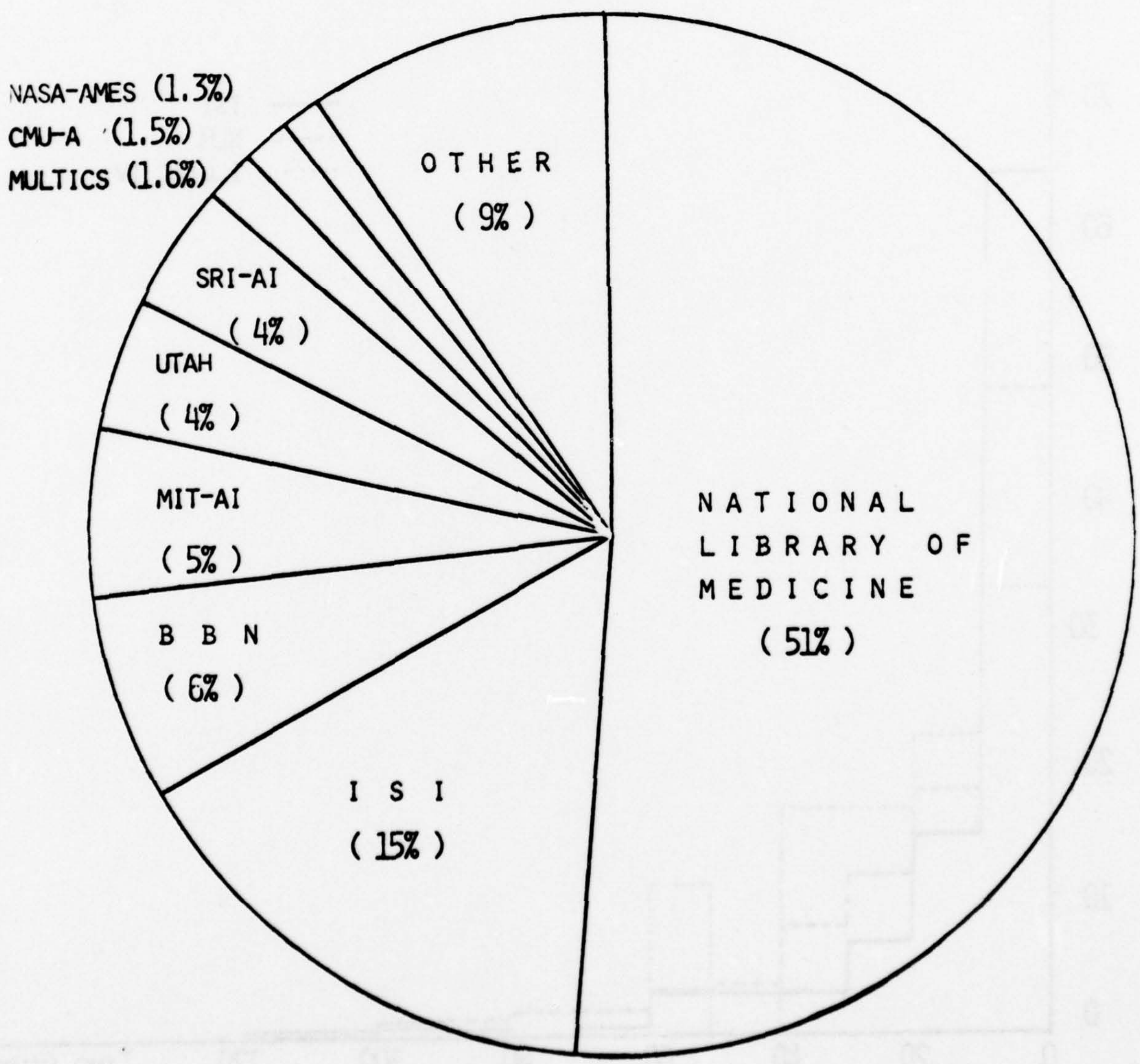
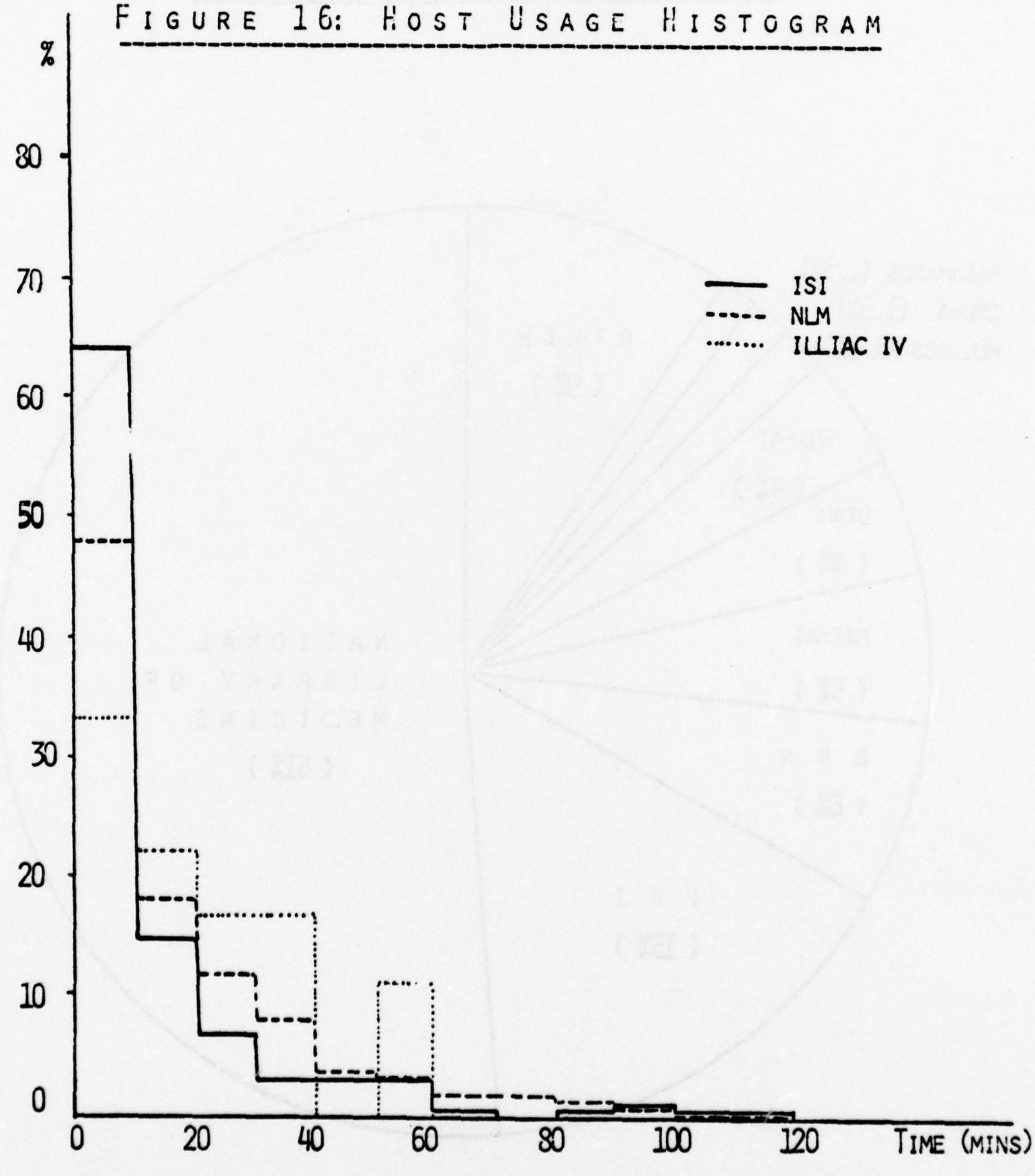


FIGURE 16: HOST USAGE HISTOGRAM



use for around two hours at a time; this seems attributable to teleconferencing and some on-line editing. The shorter periods are, in the main, due to reading and sending mail, the variations reflecting the number of messages. NLM shows a lower percentage in the 0 to 10 minute period. However, even this is probably artificially high since, due to the method of connection of NLM to ARPANET, it is possible to have an open connection to the appropriate NBS-TIP port even though NLM is unavailable. QUES records this as a connection, even though it does not appear so to the user and hence a considerable part of this 48% should be disregarded. It is not possible to say for certain how much falls into this category since there is often a need to use NLM for relatively short periods, such as reading the system news. If most of this usage is disregarded, the pattern of usage shows a maximum around the 15 minute time which is the average time for a MEDLINE search. The longer times recorded reflect multiple searches being carried out in one session, as is apparent from the more detailed data recorded by the MEDLINE program (see Sec. 2).

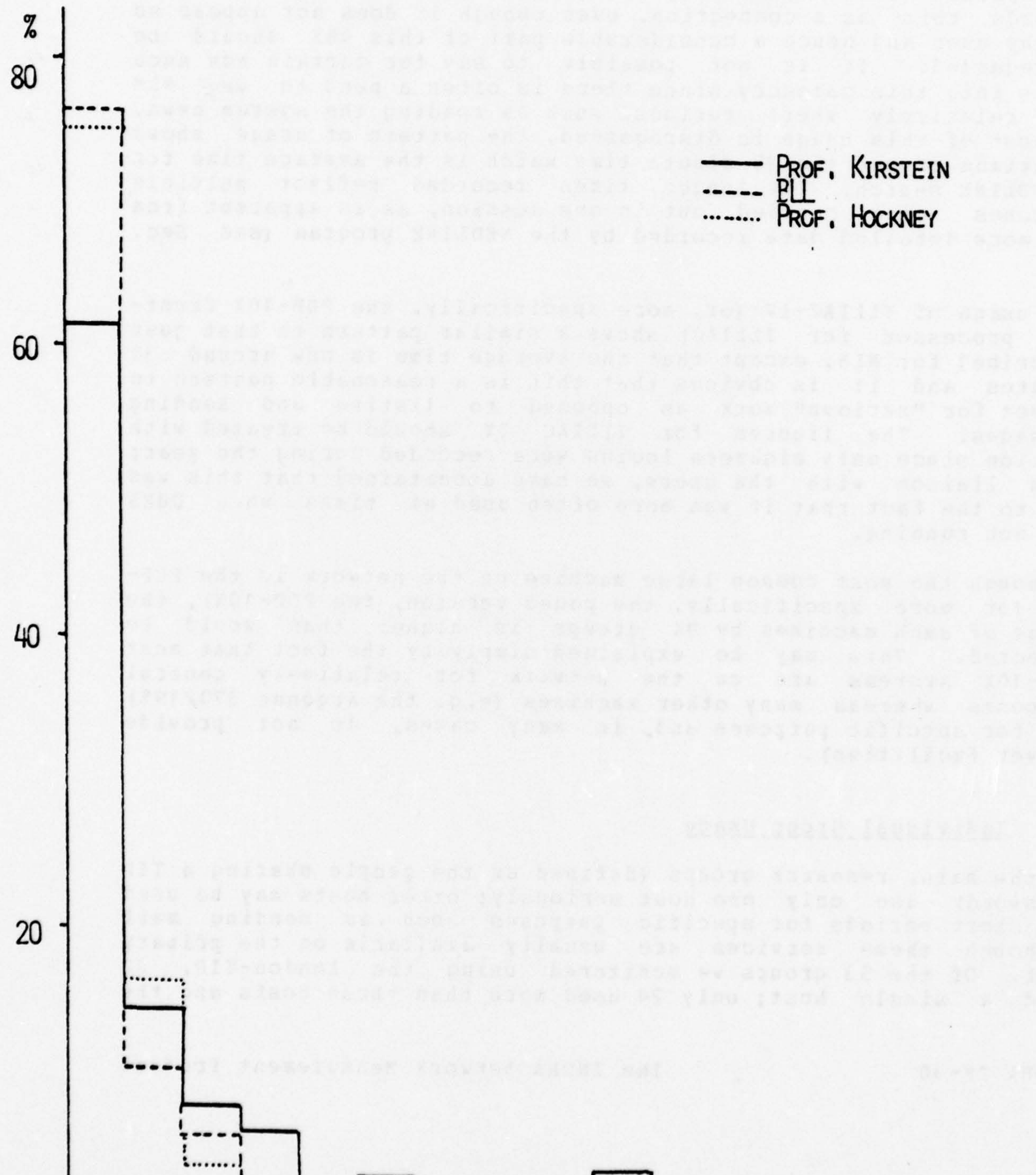
The usage of ILLIAC-IV (or, more specifically, the PDP-10X front-end processor for ILLIAC) shows a similar pattern to that just described for NLM, except that the average time is now around 30 minutes and it is obvious that this is a reasonable pattern to expect for "serious" work as opposed to listing and sending messages. The figures for ILLIAC IV should be treated with caution since only eighteen logins were recorded during the year; from liaison with the users, we have ascertained that this was due to the fact that it was more often used at times when QUES was not running.

Although the most common large machine on the network is the PDP-10 (or more specifically, the paged version, the PDP-10X), the usage of such machines by UK groups is higher than would be expected. This may be explained simply by the fact that most PDP-10X systems are on the network for relatively general purposes whereas many other machines (e.g. the Argonne 370/195) are for specific purposes and, in many cases, do not provide server facilities).

1.5 Individual Group Usage

In the main, research groups (defined as the people sharing a TIP password) use only one host seriously; other hosts may be used for short periods for specific purposes such as sending mail although these services are usually available on the primary host. Of the 53 groups we monitored using the London-TIP, 20 used a single host; only 24 used more than three hosts and the

FIGURE 17: ARPANET USAGE BY THREE RESEARCH GROUPS



highest number of hosts used was 15. Even in this last case (Thames Polytechnic), most (77%) time was used on two hosts. It must be emphasized that, at the present time, no checking is done that the user actually connects to the specified host. A visual check of the data indicates that they are accurate; any errors or deliberate falsifications are unlikely to be significant.

Due to the more or less one to one relationship between hosts and user groups, any data on group usage should correlate well with host usage. This is indeed so; however, for comparison, we show in Fig. 17 usage by Professor P.T.Kirstein (cf. Fig. 6) whose major usage is that of ISI (Host 86), Professor Hockney (Reading University), a major user of ILLIAC-IV and BLL, a major user of NLM. This matches Fig. 16 surprisingly well.

Analysis of TIP usage by user group gives an excellent check of data provided by users of the TIP (see, for example, Ref. 6) although due to some groups' usage being confined to times when QUES is rarely run, the two methods should be seen as complementary.

2. MEDLINE RESULTS

2.1 Introduction

These results may be broken down into two groups. The first are the results from the paper tape log and the second are those recorded on the magnetic tape. We consider these separately.

2.2 Global Data

These data are those recorded on the paper tape log. Use of NLM may be broken down into two distinct phases over the year. The first is use of MEDLINE alone and this experiment took place up to the end of 1975. After that, new centres took part in an experiment to evaluate the CANCERLINE database. Since there were considerable problems over the first few months of 1976, due both to hardware problems with the terminals and software problems at NLM, little access was obtained. Therefore, we confine our observations to the period July 1975 to December 1975 in this section.

As with the data reported in Part II for July, there is considerable variation in all parameters and this is clearly shown in Figs. 18 and 19 which show the usage in terms of time connected to NLM and the number of lines sent from NLM. As stated above, due to the mode of operation of QUES and the NBS-NLM interconnection, short connect times are recorded when, in fact, no real connection was made. Removing such data (and, also, unfortunately, some data where a genuine but brief connection was made), an average time per session of just under 23 minutes is found, indicating that most users perform one search per session, although, in one session, thirteen searches were performed (in nearly two hours), involving over twenty thousand characters output from NLM.

2.3 Transcript Data

In recent months, the MEDLINE program has been modified so that, if required, it is able to monitor onto either drum or magnetic tape. The data recorded is a transcript of the user's conversation with NLM with the addition of a timestamp. This data is then demultiplexed and analyzed.

At the time of writing this report, only sample results were available from this monitoring. However, we have recently obtained data for an entire week and the results from this

FIGURE 18: TIMES CONNECTED TO NLM

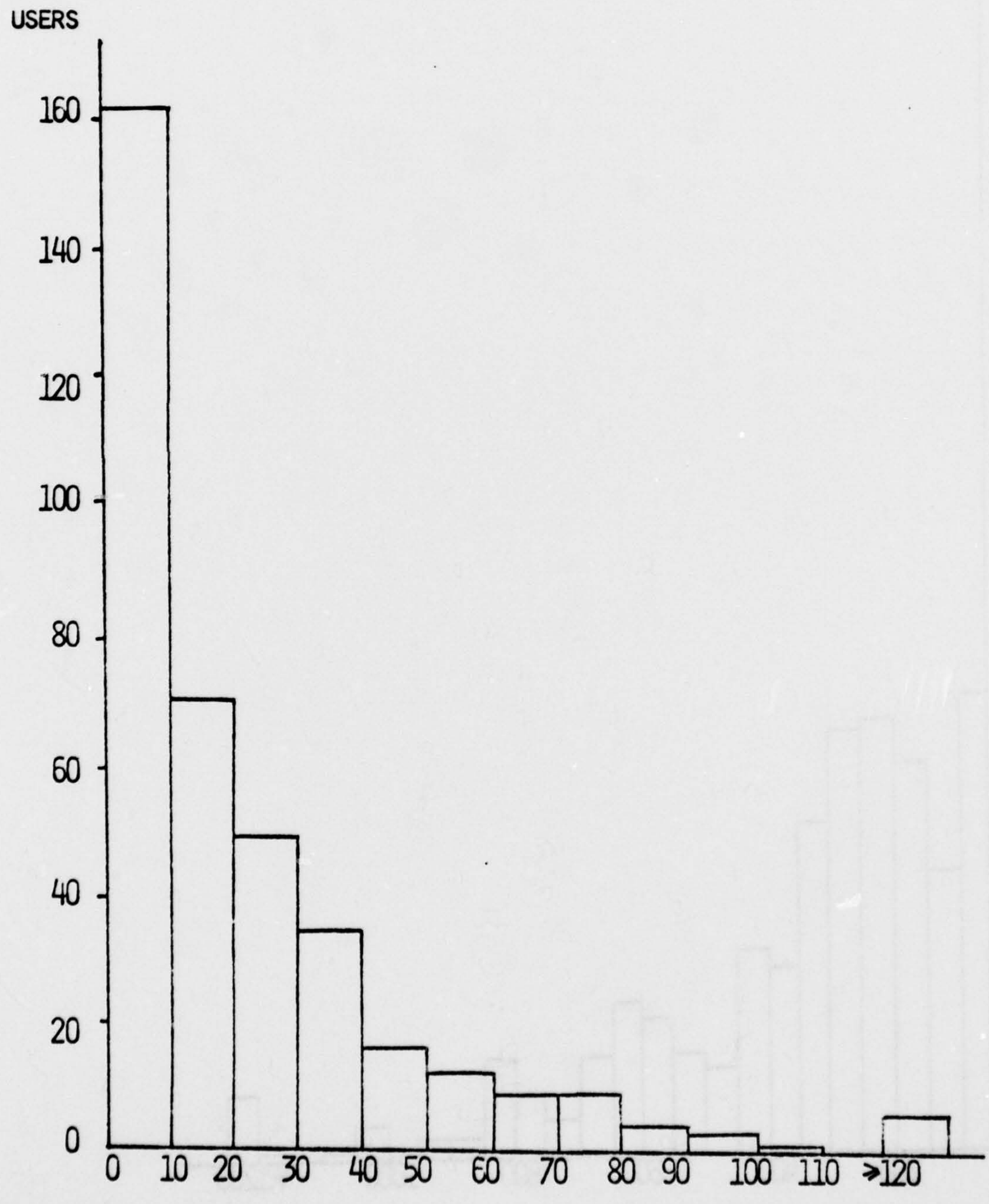
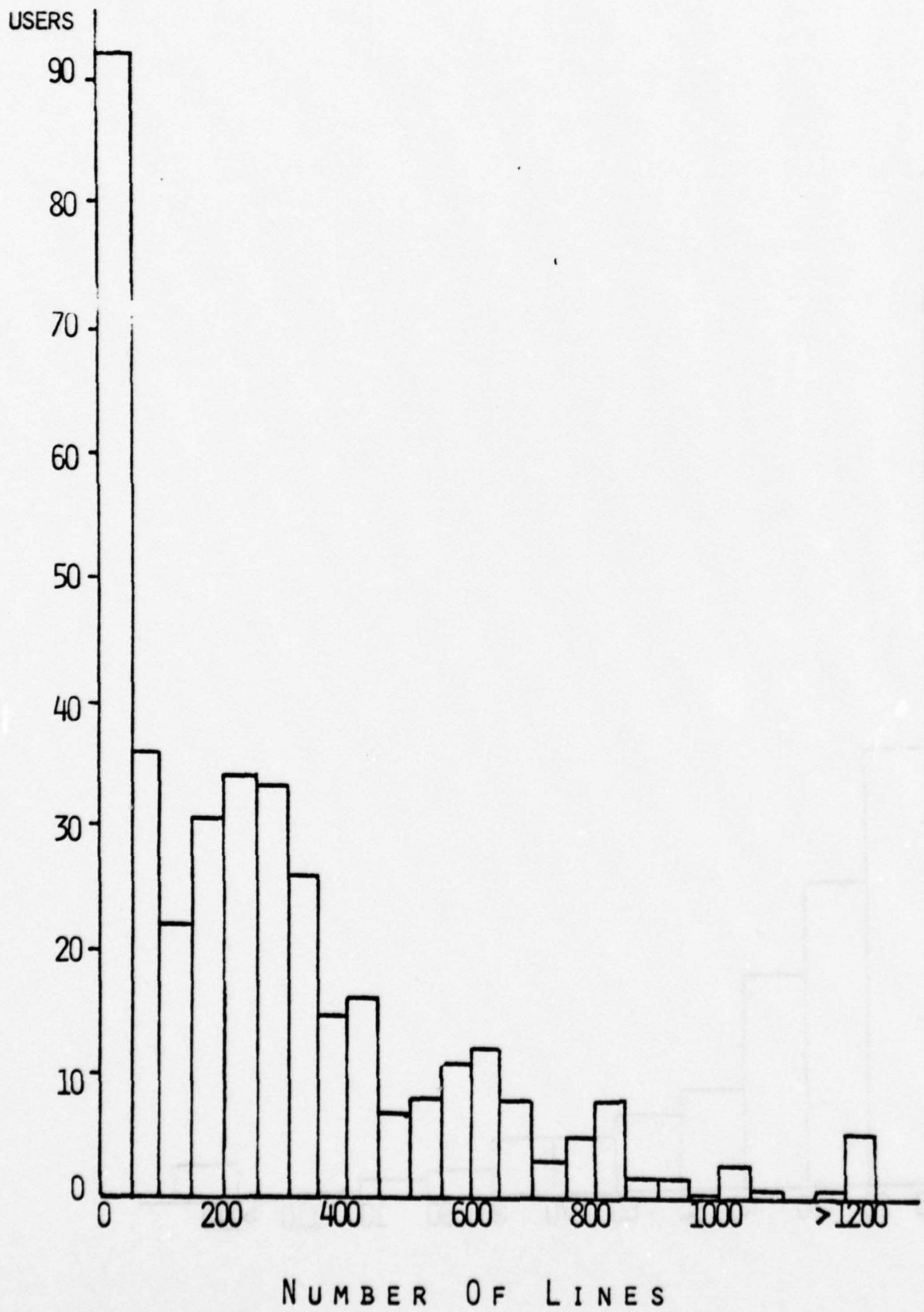


FIGURE 19: NUMBER OF LINES SENT FROM NLM



monitoring will be presented in a separate report (Ref. 24).

3. PDP-9A Results

3.1 Introduction

In all the results described up to now, we have been considering use of US hosts via ARPANET. One of the largest computers on ARPANET is the RL 360/195, accessed through PDP-9A. In a similar fashion to the PDP-9B, we capture the logging data on paper tape, send it to the RL machine and analyze it in a very similar way. These data enable us to analyze usage of one machine from various sites (i.e. the inverse of QUES monitoring) and also to examine the pattern of such usage.

In addition, we monitor similar usage to that of QUES, but instead of monitoring the TIP dial-up ports, we are monitoring users on the 360 (or one of its remote workstations) accessing ARPANET.

We present data for the year 1976. The actual analysis output is available as Ref. 23. The data gathered since the end of 1976 are being analyzed and the results will later.

3.2 Monitoring Performed and Port Usage

In principle, PDP-9A is available 24 hours a day and all monitoring is performed. In practice, this time is reduced by some (up to 2 hours a day) system development. In addition, there have been occasions when the punch has run out of tape and some data have been lost; this rarely happens and, when it does, it is usually on Sunday evening during which time there is little activity. Nevertheless, the figures quoted are a slight underestimate due to this. The actual amount of monitoring performed and for which data are available is shown in Figs 20 and 21, the former being broken down by month, the latter by hour of day. It is clear from the figure by month that little data was available for January and not all for October. In addition, it shows the effect of the move from Gordon Square to Gower Street in July and the fact that the machine was switched off for a week over the Christmas period. The graph by hour of day shows a remarkable consistency, indicating merely that the machine was used for system development in the afternoon and early evenings, thus reducing its availability in those periods.

The first figures we describe are those of the computer up-times. PDP-9A was available for nearly 60% of the time but this figure is probably a considerable underestimate as some data are not available. It was deliberately shut down 100 times and crashed

FIGURE 20: PDP-9A MONITORING PERFORMED BY MONTH

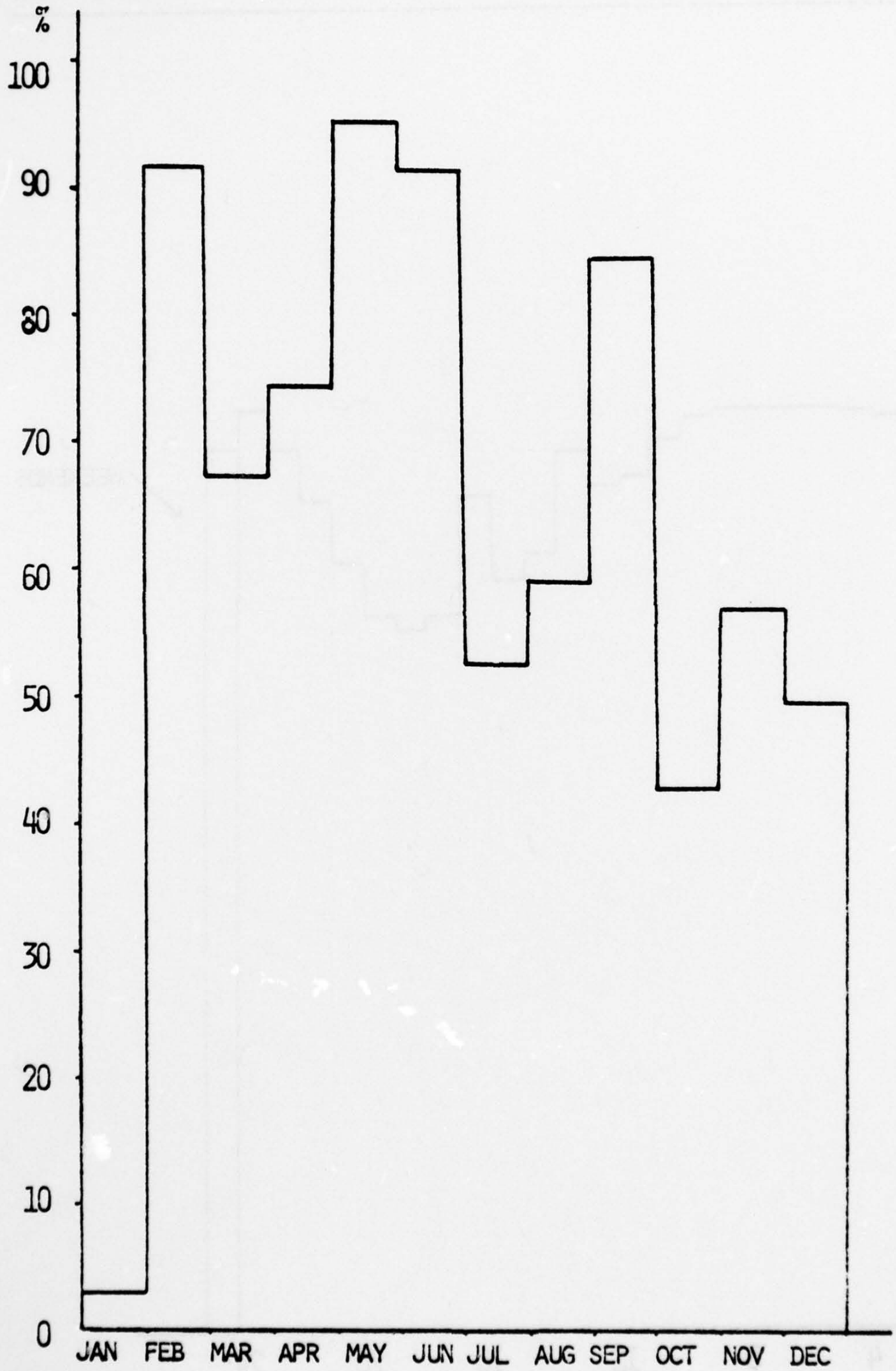
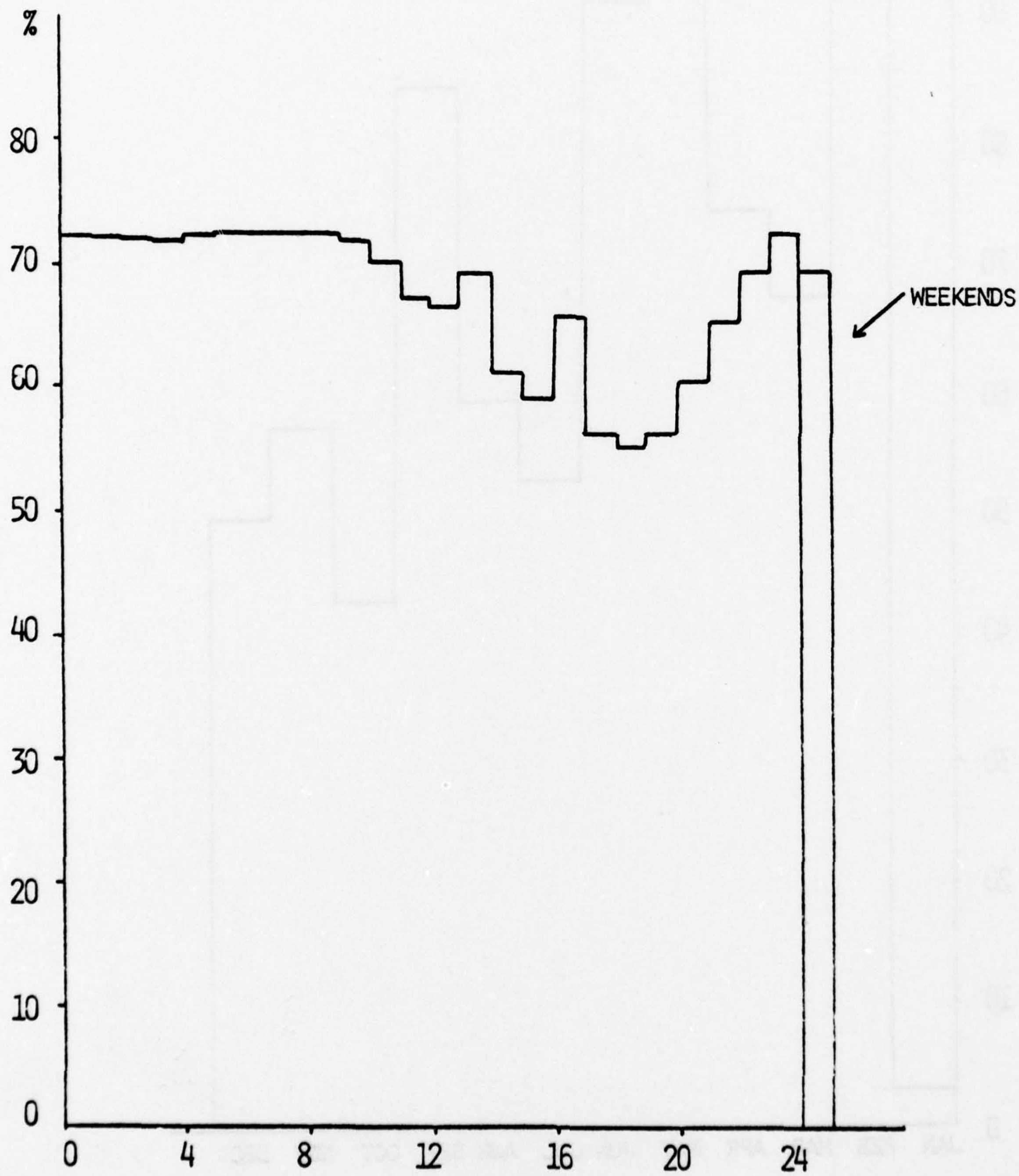


FIGURE 21: PDP-9A MONITORING PERFORMED AS A FUNCTION OF TIME OF DAY



847 times (although if it is shut down incorrectly, it may be recorded as a crash) giving an MTBF of over six hours. However, on five occasions, a serious fault developed which caused the system to crash repeatedly (since, on a system crash, there is an autorestart mechanism). These five occasions gave rise to a total of 307 crashes and hence a more realistic figure of the number of crashes would be 545 giving a true MTBF of over 9 hours. Obviously all the other figures for up times must be lower than that for the PDP-9 and the figures we quote are absolute with the relative figures in brackets. The IMP was by far the most reliable with an up time of 55% (95%) and only down 136 times. RL was available 48% (83%) and RSRE 23% (39%).

We also measure port usage which was fairly heavy. It is not possible from the statistics recorded to determine how many attempts to connect to the PDP-9 were rejected due to all ports being in use. However, in the time considered, all ports were occupied 155 times for a total of 174 minutes.

In the case of the RL 360, it is possible to determine the same data (422 times and 1165 minutes), as well as the number of times (358) when a user had logged into the PDP-9 and was then unable to log into the 360 due to all the ports being in use. This is not a particularly valuable datum since a user may attempt to log in a number of times in a brief time interval. Nevertheless, it is significant that so many attempts were made and this is borne out by comments from users as to their inability to access RL at various times of day, particularly in late afternoon London time.

This is shown clearly in Figs. 22 and 23, the former relating to the PDP-9, the latter to RL. The number of logins per hour follows a similar pattern and is omitted from these graphs for clarity. It can be seen that the port usage of the two machines is, as expected, virtually identical. The peak period is around 1500 to 1600 local (BST or GMT as applicable) when heavy usage from local users coincides with early morning usage from the West Coast of the USA and midday usage from the East Coast.

3.3 Connection Statistics

Perhaps the most significant item in these results is the large number of 360 IDs used (96) and the pattern of usage which differs noticeably from that recorded by QUES described above. In particular, the average times are much higher (an overall average of 27 minutes). On examining the statistics grouped by ident, it is noticeable that usage by the INDRA group has a larger average connect time than other groups. In these statistics, we have not differentiated between local use and

FIGURE 22: PDP-9A PORT USAGE BY TIME OF DAY

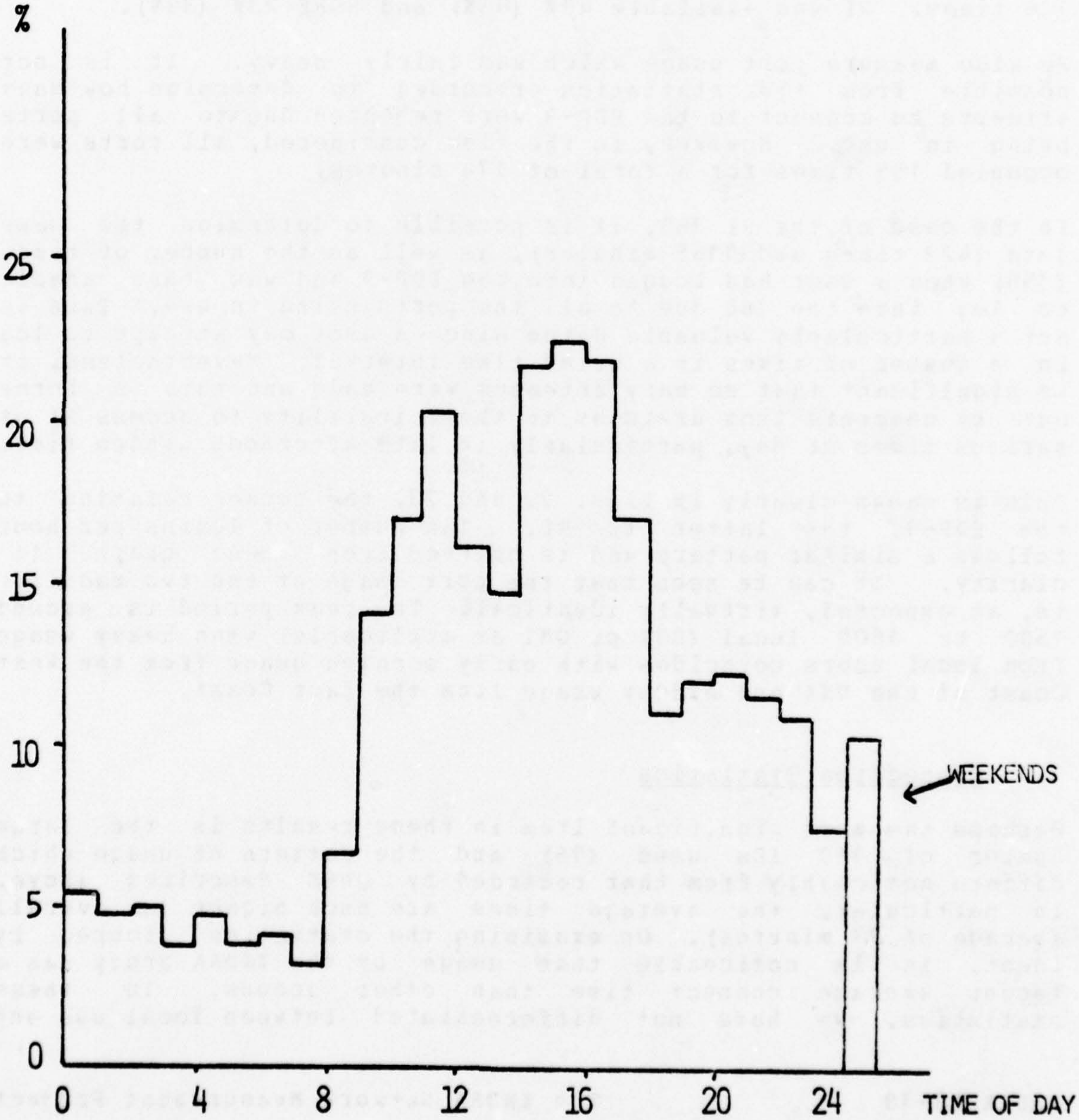
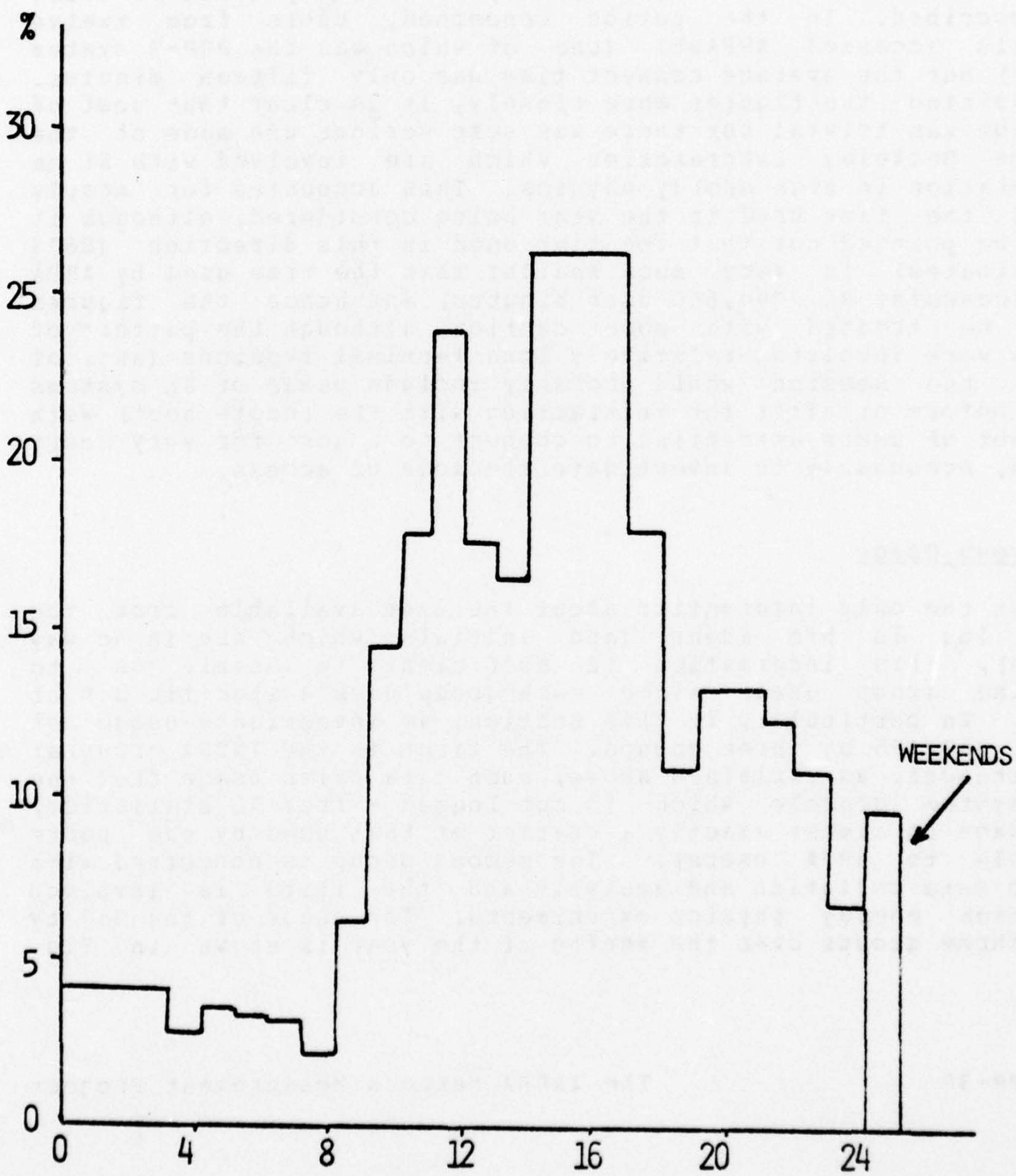


FIGURE 23: RL PORT USAGE BY TIME OF DAY



remote use (which could be done by examining the TIP port number) but it is certain that the great majority of INDRA group usage is via hard-wired terminals and this may well influence usage habits.

As would be expected, such usage is the highest made of the RL 360 and this is shown diagrammatically in Fig. 24. As would be expected, this usage is mirrored by the pattern of hosts used to access RL with the preponderance of time used (nearly three-quarters) being from the London-TIP as shown in Fig. 25.

We are also able to examine usage from terminals attached to the RL 360 and these show a considerably different pattern to those just described. In the period concerned, users from twelve terminals accessed ARPANET (one of which was the PDP-9 system console) but the average connect time was only fifteen minutes. On examining the figures more closely, it is clear that most of the usage was trivial but there was some serious use made of the Lawrence Berkeley Laboratories which are involved with RL on collaboration in high energy physics. This accounted for nearly 90% of the time used in the year being considered, although it should be pointed out that the time used in this direction (8611 user minutes) is very much smaller than the time used by ARPA users accessing RL (144,690 user minutes) and hence the figures should be treated with some caution, although the pattern of serious work involving relatively long terminal sessions (and, of course, the session would probably include usage of RL systems either before or after the interaction with the remote host) with a number of users attempting to connect to a host for very short periods, presumably to investigate the mode of access.

3.4 Group Usage

Although the only information about the user available from the PDP-9A log is his ident (and initials which are in no way checked), this information is sufficient to enable us to determine group usage since each group uses a specific set of idents. In particular, in this section, we investigate usage of the RL 360/195 by three groups. The first is the INDRA group at UCL (although, as explained above, such data omits usage from the HASP system console which is not logged - from RL statistics, this usage is almost exactly a quarter of that used by the ports available to ARPA users). The second group is concerned with seismic data collation and analysis and the third is involved with high energy physics experiments. The usage of the 360 by these three groups over the months of the year is shown in Fig. 26.

FIGURE 24: USAGE OF RL 360/195 BY VARIOUS RESEARCH PROJECTS

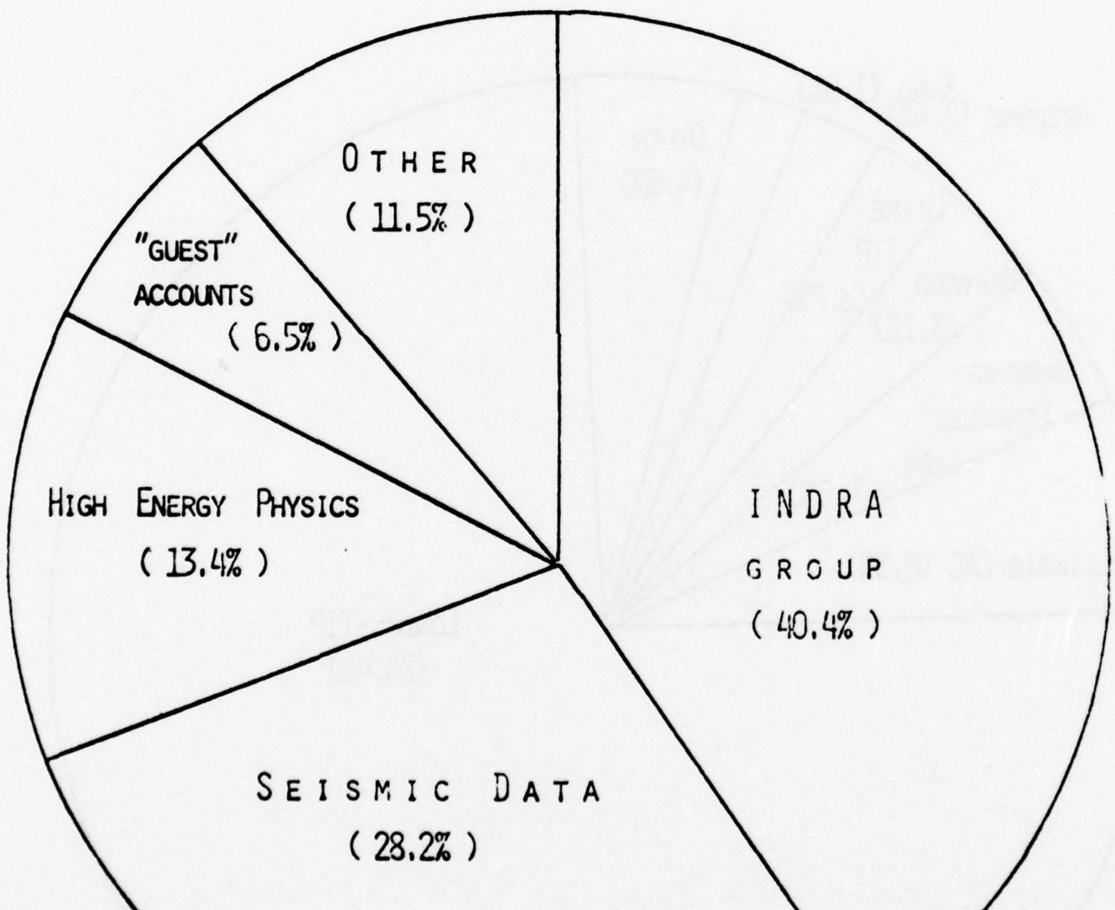


FIGURE 25: USAGE OF RL 360/195 FROM VARIOUS ARPANET HOSTS

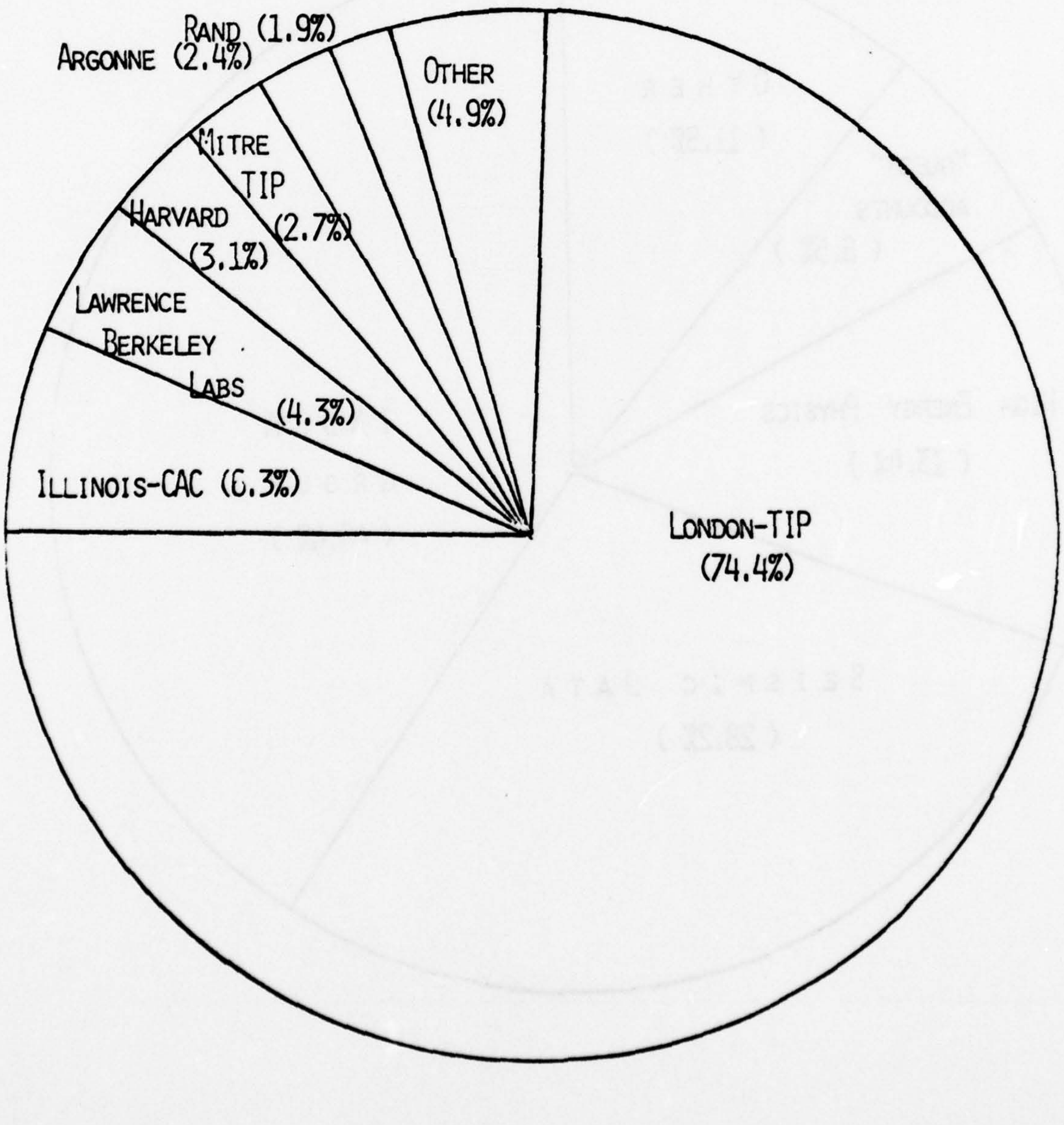
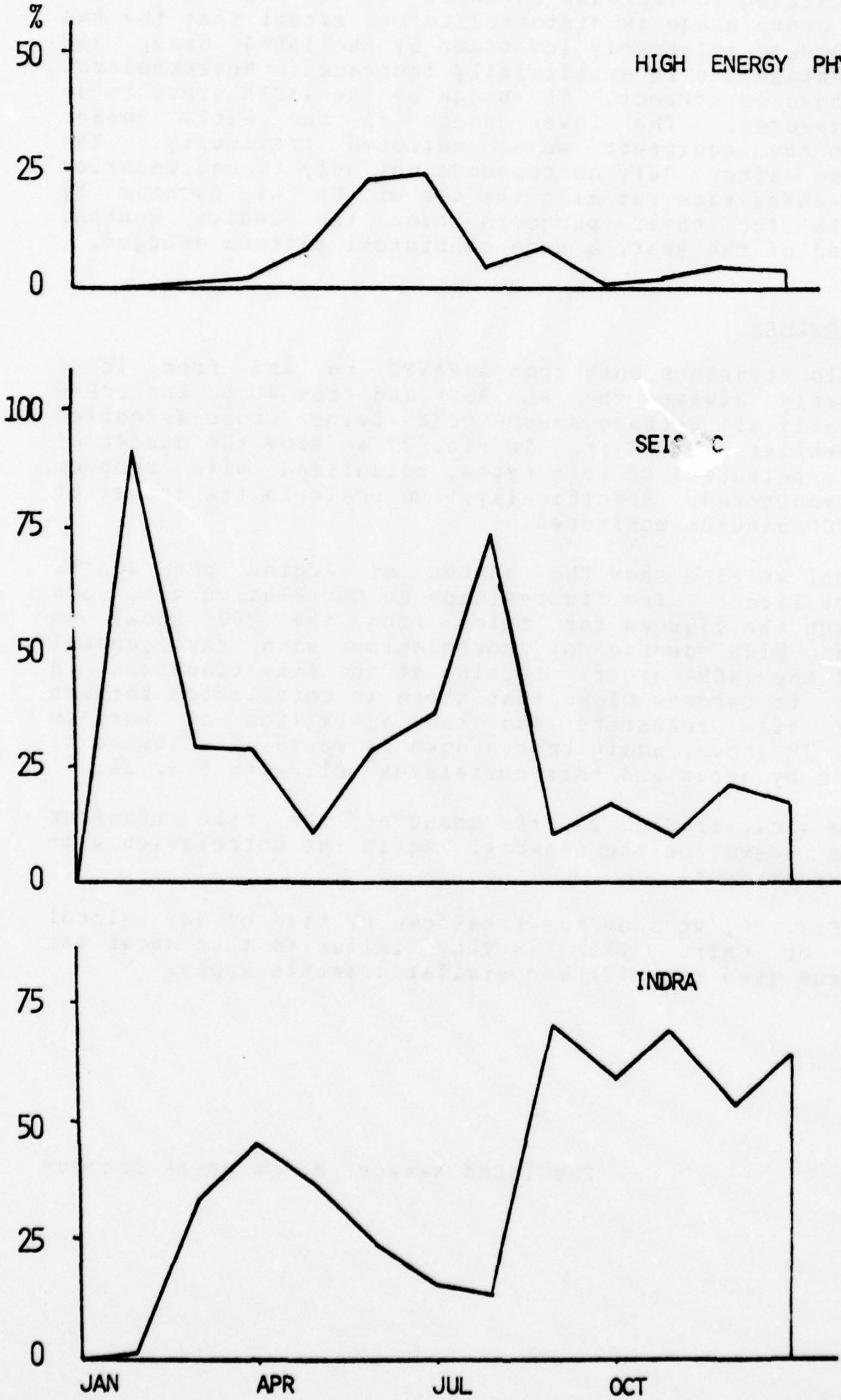


FIGURE 26: USAGE OF RL 360/195 BY THREE RESEARCH GROUPS



The ordinate shows the percentage of total use of the PDP-9 for the specified month. It is clear that the HEP groups used only a relatively low proportion of the time, rising from nothing at the beginning of the year to a peak in June, tailing off towards the autumn then starting to increase again at the end of the year. The seismic group usage is distorted to the extent that the two peaks correspond to relatively low usage by the INDRA group and thus the percentage use is artificially increased. Nevertheless, the pattern shown is correct. The usage by the INDRA group is as would be expected. The lower usage in the early summer corresponds to the equipment move mentioned previously. The rapid increase after July corresponds not only to compensation for lack of machine time but also the use of the RI machine by M.Sc. students for their projects over the summer months. Towards the end of the year, a more consistent pattern emerged.

3.5 File Transfers

We measure file transfers both from ARPANET to and from local machines (nearly always the RL 360) and from RL to the PDP-9 (which are nearly all Babbage source code being cross-Assembled to PDP-9 Assembler (Ref. 17)). In Fig. 27 we show the number of file transfers performed of both types, normalized with respect to the time monitored. Specifically, the scale is the number of events per 1000 minutes monitored.

For comparison, we also show the number of logins per month, suitably normalized. The figures show no correlation with each other, although the figures for files from the 360 show, as expected, a high degree of correlation with the general activities of the INDRA group. Looking at the file transfers in more detail, it becomes clear that there is correlation between the number of file transfers and the activities of various groups. Fig. 28 shows, again broken down by month, the number of file transfers by group and this correlates well with Fig. 26.

Similarly, we show, in Fig. 29, the usage of the file transfer from various hosts on the network. Again the correlation with group usage is evident.

Finally, in Fig. 30, we show the breakdown by time of day (local time - BST or GMT). This is very similar to that shown for terminal access (see Fig. 12) and similar comments apply.

FIGURE 27: FILE TRANSFERS BY MONTH

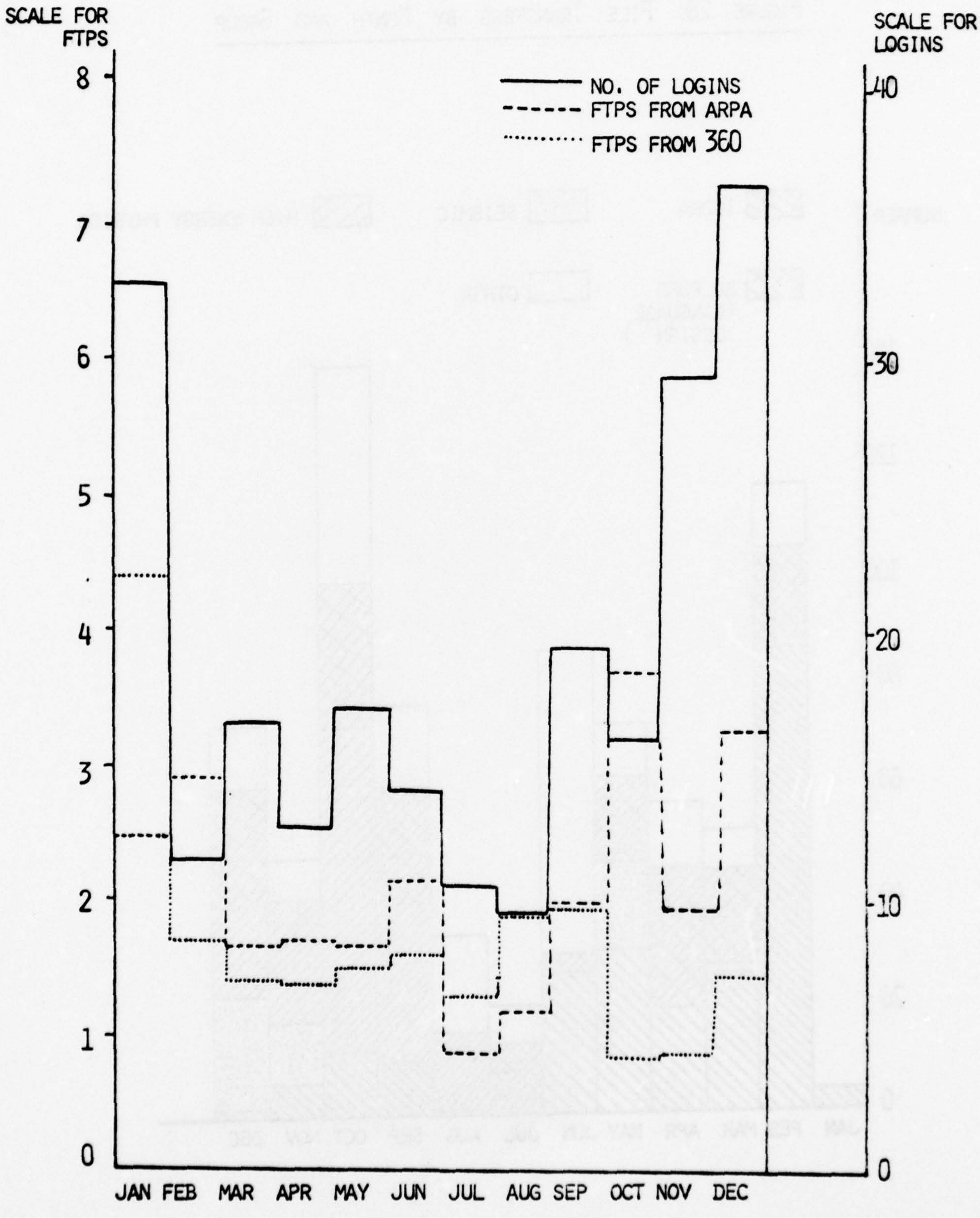


FIGURE 28: FILE TRANSFERS BY MONTH AND GROUP

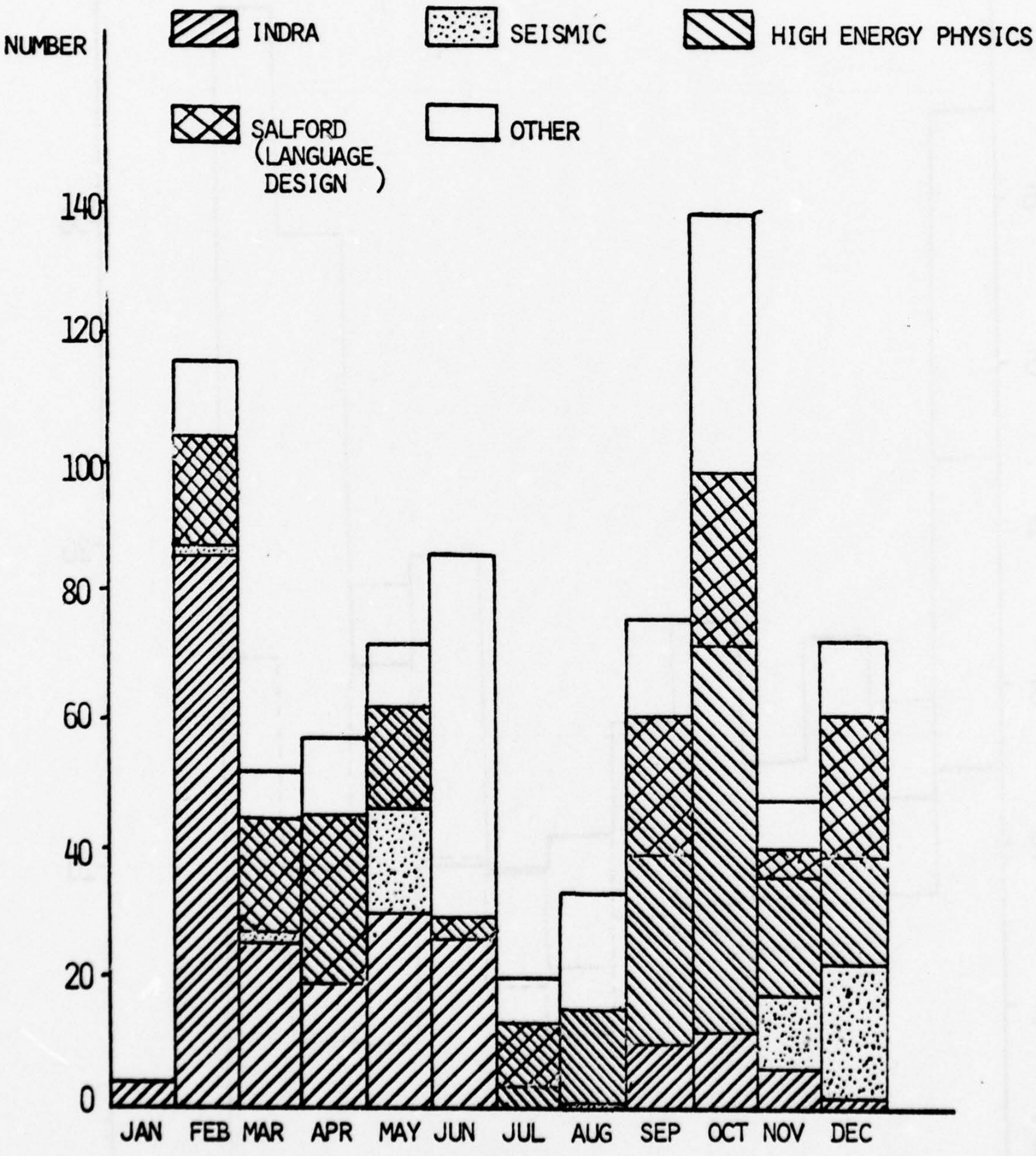


FIGURE 29: USAGE OF ARPANET HOSTS FOR FILE TRANSFERS

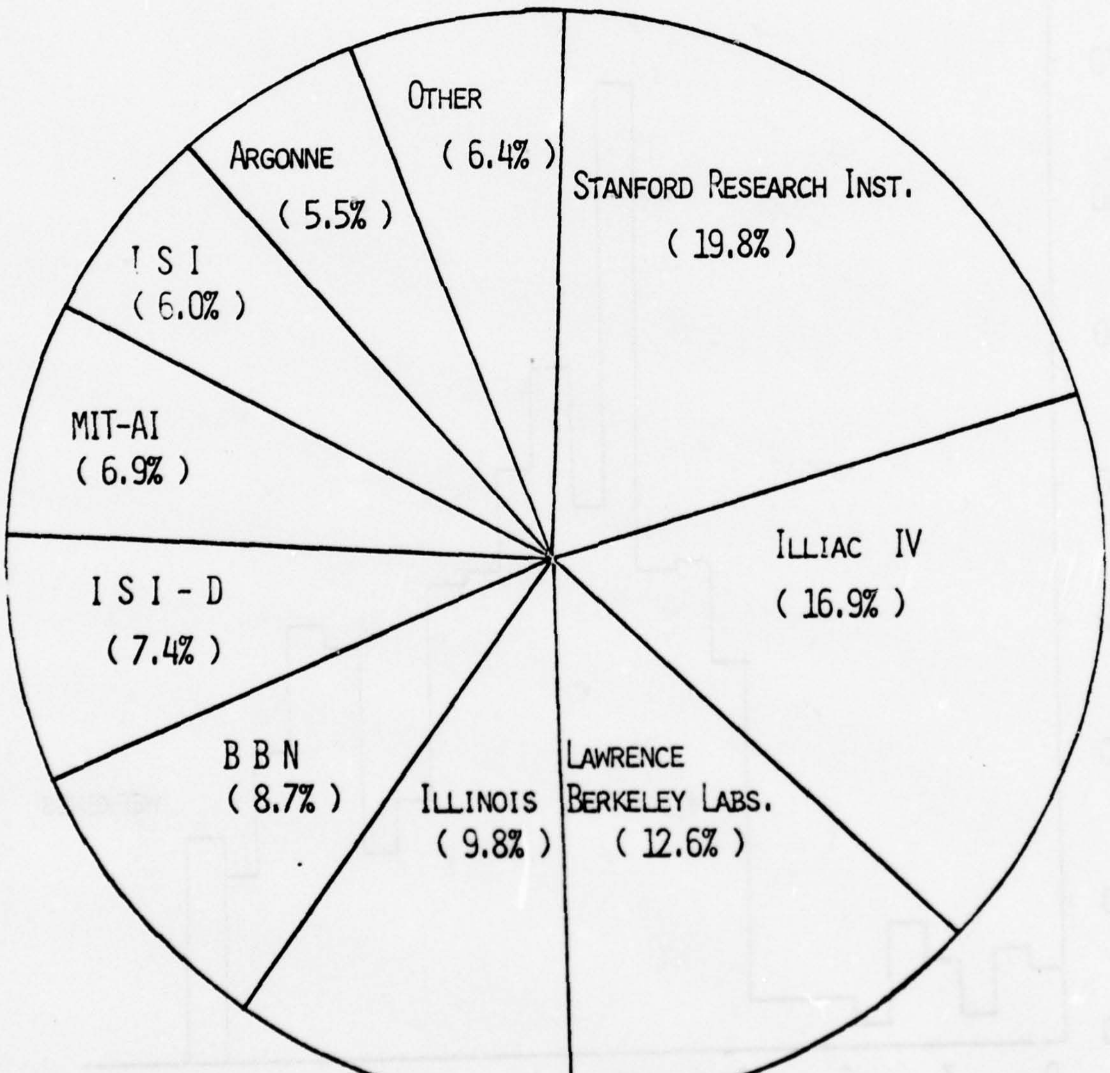
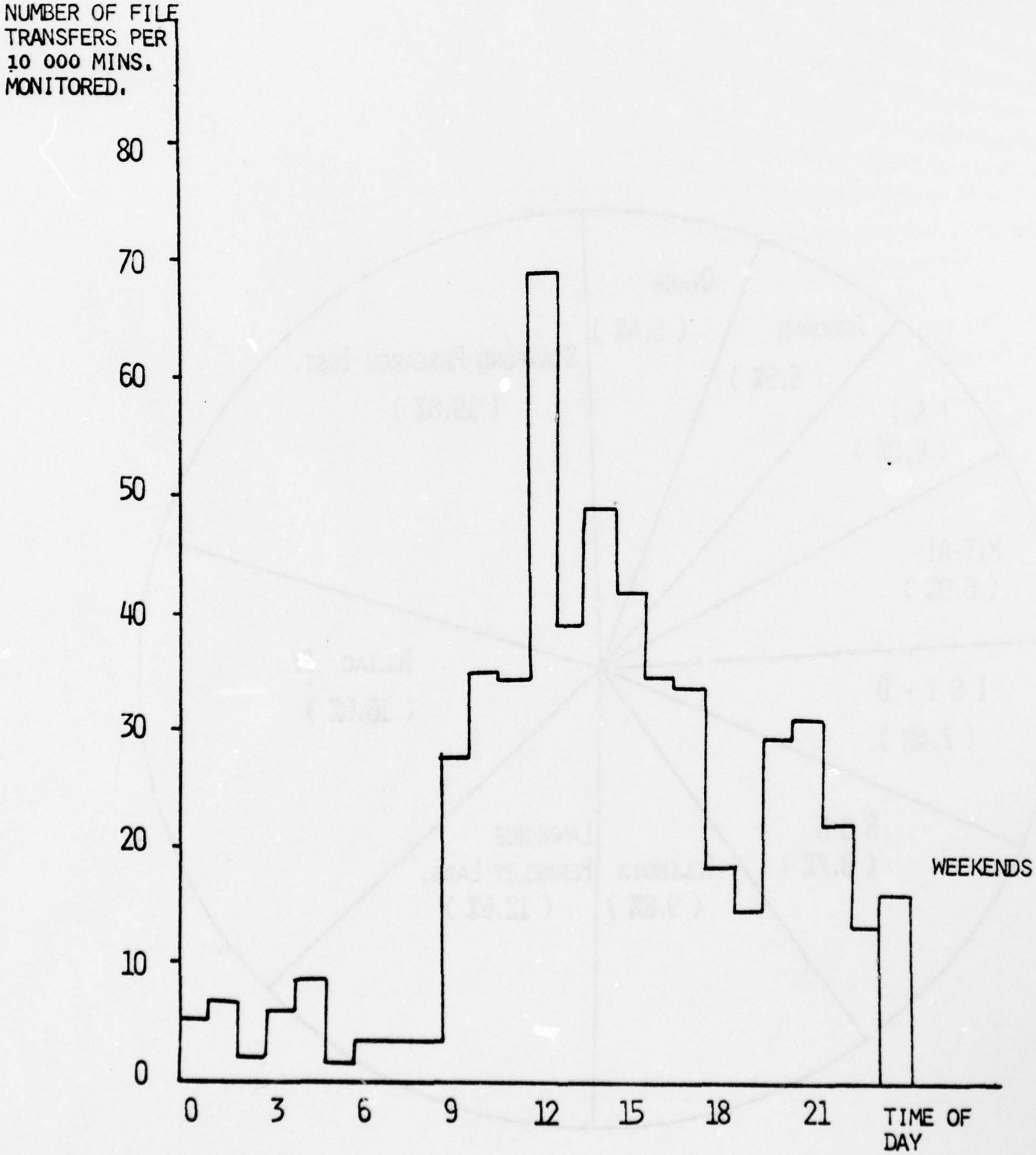


FIGURE 30: FILE TRANSFERS BY TIME OF DAY



PART IV

A USER GUIDE

1. PDP-9 PROGRAMS

1.1 QUES/MEDLINE

This is the standard system to run on PDP-9(B). To run it, all that is necessary is to bootstrap the system from tape 151. Various ACC switches may be set with the following effects:

1	Auto restart in the event of a system crash
3	Do not monitor MEDLINE program
6	Monitor to drum rather than tape
7	Trace system variables

A tape must be available on the tape drive with write enabled, unless either switch 3 or switch 6 is set.

2. RL 360/195 PROGRAMS

2.1 XFSTATS

A user guide to the XFSTATS package, including some details of the programs described in detail above is available as Ref. 20, as is a description of the procedure, also called XFSTATS.

XFSTATS is a member of the ELECTRIC public library JOBFILE and is accessed by:

EXEC JB=XFSTATS,.....

where all the other parameters are, due to the method of operation of ELECTRIC (Ref. 3), keyword parameters. The allowed keywords, their defaults and meanings are shown below:

Name	Default	Meaning
PRI	12	Priority
ACCT	Logged in A/C	Account
ID	Logged in ID	Identifier
LINES	--	Lines of output
TIME	--	CPU time required
COPIES	1	Copies of output
LPARM	NOMAP,NOLIST	Link-edit parameters
V	-	Version of program (null or X)
FROM	1/7/75	Start date
TO	1/7/76	Finish date
OP	--	Operation
OP2	--	"
OP3	--	"
OP4	--	"
OP5	--	"
CONTROL	--	File of control cards
DUMP	Lineprinter	Dump stream

The global edit (EL) is provided to enable the output to be written back into Electric. If this is used, the following plants are allowed:

SPP	6	Primary space (tracks)
SPS	5	Secondary space (tracks)
VOL	RHEL04	Disk volume for temporary output
KEY	AVS	Key of Electric user

Global edits also exist for producing output on 35mm film or

microfiche, as well as for the use of the RL graphical routines (ENPLOT).

PART V

CONCLUSIONS

1. METHODS

The method of monitoring and access control (QUES) we have developed is not of general application, particularly since it requires a dedicated processor in addition to the one being monitored and, in the general case, it is obviously simpler to put these functions in the latter. However, the introduction of monitoring and/or access controls into a computer brings with it a loss of reliability (since, not only does it increase the complexity of the software but also it requires additional hardware). This loss of reliability may not be acceptable and so our technique may be of more general applicability.

An example of this is evinced by ARPANET. Access controls were introduced into the TIPs in late 1974 and, for a number of reasons (in particular, the decrease of reliability that backing store would introduce and the problems of maintaining the password data base) it was introduced in a way that led to inconvenience for users and large network overheads. Although our approach would not obviate the problem of maintaining the data base, it would certainly not decrease the reliability of the TIPs. At the present time, the LONDON-TIP is the only node on ARPANET on which monitoring of users and access control is being carried out.

The technique of recording the data onto paper tape has certain advantages, in particular the fact that a system crash cannot corrupt the data. However, it has many attendant disadvantages and a more mechanical method of sending the recorded data to the machine on which it is to be analyzed would be of considerable value. This may be done with reasonably little effort on PDP-9A but, since PDP-9B is not connected to the RL machine, it would prove more difficult in that case. Various problems in the analysis have been found and this experience will enable us to overcome such problems in future. For example, a major problem with analyzing the PDP-9A data has been the fact that users frequently give the machine the wrong date and/or time and this means that the data has to be edited manually. It would be fairly simple to check the time with that on a connected machine and, perhaps, to flag an error if a discrepancy were noted.

2. NETWORK USAGE

Using the techniques we have described, we have been able to obtain a fairly accurate picture of the extent of usage of the network for different applications. The measurements give an excellent cross-check on subjective reports from user groups on their usage of specific Hosts (Ref. 7).

The monitoring performed on PDP-9A is of similar interest, giving us a picture both of the usage of one large machine on ARPANET and of the usage of ARPANET through such a machine. Much of the data obtained correlates that expected or experienced qualitatively but nevertheless, a quantitative evaluation is valuable. A very large volume of data has been collected and analyzed and the results are presented in detail in this report. However, a number of conclusions about network usage stand out. These are, in no particular order:

(i) Sessions from hardwired terminals are longer than those through the PSTN

(ii) The most frequently used computer system on the network is the Tenex (even more than would be expected from the predominance of such machines)

(iii) Far greater use is made of the RL 360 from ARPANET than the converse.

3. OVERALL CONCLUSIONS

We have developed a number of powerful access control and monitoring techniques, including both data acquisition and analysis. These methods have proved of great value and complement other monitoring performed on ARPANET.

Use of these tools has enabled us to obtain a fairly complete picture of the usage of the UC node of ARPANET both via the PSTN and directly. We have been able to verify users subjective reports and perform quantitative measurements on many aspects of usage.

Thus our techniques have proved of considerable use in the past and can continue to do so in future.

APPENDICES

APPENDIX I: SAMPLE DATA

These data are intended to show the input and output of the various analysis programs. In order to fit into the page format, some slight editing has been done.

A1.1 Sample PDP-9B Raw Data

```
14 JUL 75 12 45 QUE:1 **QUES MONITORING 6 PORTS FROM PORT# 70
14 JUL 75 12 46 MED:5 TIME MCH MLNS UCH ULNS @S "STO "PRI
14 JUL 75 13 05 QUE:2 UNKNOWN HOST NUMBER
14 JUL 75 13 05 QUE:724         QMC 87
14 JUL 75 13 05 QUE:724         QMC 86
14 JUL 75 13 10 QUE:74FORSEY    WESS147
14 JUL 75 13 10 MED:74FORSEY    WESS147
14 JUL 75 13 13 MED:74  000106  0147 008   027 003   00   00   00
14 JUL 75 13 13 QUE:74000105
14 JUL 75 13 14 QUE:75FORSEY    WESS147
14 JUL 75 13 14 MED:75FORSEY    WESS147
14 JUL 75 13 18 MED:75  000130  0000 011   036 004   00   00   00
14 JUL 75 13 18 QUE:75000129
14 JUL 75 13 18 QUE:2 UNKNOWN HOST NUMBER
14 JUL 75 13 18 QUE:75FORSEY    WESS146
14 JUL 75 13 19 QUE:75FORSEY    WESS147
14 JUL 75 13 19 MED:75FORSEY    WESS147
14 JUL 75 13 21 QUE:6 QUES OK
14 JUL 75 13 21 MED:75  000033  0000 003   009 001   00   00   00
14 JUL 75 13 21 QUE:75000032
14 JUL 75 13 21 QUE:2 UNKNOWN HOST NUMBER
14 JUL 75 13 21 QUE:75FORSEY    WESS%HME
14 JUL 75 13 21 QUE:2 UNKNOWN HOST NUMBER...ACCESS PROHIBITED
14 JUL 75 13 21 QUE:75FORSEY    WESS%HME
14 JUL 75 13 52 QUE:72003810
14 JUL 75 13 52 QUE:724         QMC 66
14 JUL 75 13 57 QUE:6 QUES OK
14 JUL 75 14 13 QUE:72001652
14 JUL 75 14 18 QUE:8 MONITORING TERMINATED
```

A1.2 Sample PDP-9P Packed Data

```
14 07 75 12 45 Q#1
14 07 75 13 05 Q72 4          CMC  86  00 38 10
14 07 75 13 05 Q#9 03
14 07 75 13 10 Q74 FORSEY     WESS 147 00 01 05
14 07 75 13 10 Q#9 04
14 07 75 13 13 Q#9 03
14 07 75 13 14 Q75 FORSEY     WESS 147 00 01 29
14 07 75 13 14 Q#9 04
14 07 75 13 18 Q#9 03
14 07 75 13 19 Q75 FORSEY     WESS 147 00 00 32
14 07 75 13 19 Q#9 04
14 07 75 13 21 Q#9 03
14 07 75 13 52 Q#9 02
14 07 75 13 52 Q72 4          CMC  66  00 16 52
14 07 75 13 52 Q#9 03
14 07 75 14 13 Q#9 02
14 07 75 14 18 Q#8
```

A1.3 Sample_MEDLINE_Raw_Data

```
05 MAY 76 02 58 03.1 *<MEDLINE MONITORING TO DRUM
05 MAY 76 02 59 01.0 *<1  MEDLINE PROGRAM VERSION 2.06
05 MAY 76 09 20 33.5 *<70ROBERTS ICRF147
05 MAY 76 09 20 34.1 *<70ROBERTS ICRF147
05 MAY 76 09 20 40.7 K<WHEN I LAST LOOKED A FEW MINUTES AGO, $
05 MAY 76 09 20 41.1 K<MEDLINE WAS AVAILABLE
05 MAY 76 09 20 42.8 K<TERMINAL (FAST OR SLOW) >$
05 MAY 76 09 20 51.0 K>FAST
05 MAY 76 09 20 53.2 K<PLEASE WAIT WHILE I TRY TO CONNECT YOU...
05 MAY 76 09 20 54.1 *<TRYING PORT #144
05 MAY 76 09 20 54.5 K<TRYING PORT #144
05 MAY 76 09 20 59.2 *<TRYING PORT #142
05 MAY 76 09 20 59.7 K<TRYING PORT #142
05 MAY 76 09 21 17.3 *<TRYING PORT #140
05 MAY 76 09 21 17.7 K<TRYING PORT #140
05 MAY 76 09 21 39.2 *<TRYING PORT #136
05 MAY 76 09 21 39.6 K<TRYING PORT #136
05 MAY 76 09 21 48.3 *<TRYING PORT #134
05 MAY 76 09 21 48.8 K<TRYING PORT #134
05 MAY 76 09 21 51.0 *<STATUS CODE OF MEDLINE IS 08
05 MAY 76 09 21 51.8 K<THERE WAS NET TROUBLE
05 MAY 76 09 21 52.9 K<PLEASE TRY AGAIN SHORTLY - BYE FOR NOW
05 MAY 76 09 30 01.0 *<72ROBD     ICRF147
05 MAY 76 09 30 01.7 *<72ROBD     ICRF147
05 MAY 76 09 30 05.3 I<WHEN I LAST LOOKED A FEW MINUTES AGO, $
05 MAY 76 09 30 05.7 I<THERE WAS NET TROUBLE
05 MAY 76 09 30 06.4 I<SHALL I TRY AGAIN (YES OR NO) > $
05 MAY 76 09 30 10.0 I>YES
05 MAY 76 09 30 12.1 I<TERMINAL (FAST OR SLOW) >$
05 MAY 76 09 30 15.3 I>FAST
05 MAY 76 09 30 17.1 I<PLEASE WAIT WHILE I TRY TO CONNECT YOU...
05 MAY 76 09 30 18.3 *<TRYING PORT #144
05 MAY 76 09 30 18.7 I<TRYING PORT #144
05 MAY 76 09 30 44.9 *<TRYING PORT #142
05 MAY 76 09 30 45.3 I<TRYING PORT #142
05 MAY 76 09 30 52.1 *<TRYING PORT #140
05 MAY 76 09 30 52.6 I<TRYING PORT #140
05 MAY 76 09 31 10.3 *<TRYING PORT #136
05 MAY 76 09 31 10.7 I<TRYING PORT #136
05 MAY 76 09 31 29.8 *<TRYING PORT #134
05 MAY 76 09 31 30.2 I<TRYING PORT #134
05 MAY 76 09 31 50.3 *<STATUS CODE OF MEDLINE IS 08
05 MAY 76 09 31 54.6 I<THERE WAS NET TROUBLE
05 MAY 76 09 31 55.0 I<PLEASE TRY AGAIN SHORTLY - BYE FOR NOW
05 MAY 76 09 33 03.2 *<8  MEDLINE MONITORING TERMINATED
```

A1.4 Sample MEDLINE Demultiplexed Output

User ROBERTS from ICRF on TIP port #70

Started 05 MAY 76 09 20

```
05 05 76 09 20 40.7: WHEN ... MINUTES AGO, MEDLINE WAS AVAILABLE
05 05 76 09 20 42.8: TERMINAL (FAST CR SLOW) >FAST
05 05 76 09 20 53.2: PLEASE WAIT WHILE I TRY TO CONNECT YOU...
05 05 76 09 20 54.1: TRYING PORT #144
05 05 76 09 20 59.2: TRYING PORT #142
05 05 76 09 21 17.3: TRYING PORT #140
05 05 76 09 21 39.2: TRYING PORT #136
05 05 76 09 21 48.3: TRYING PORT #134
05 05 76 09 21 51.8: THERE WAS NET TROUBLE
05 05 76 09 21 52.9: PLEASE TRY AGAIN SHORTLY - BYE FOR NOW
```

A1.5 Sample CUPS Analysis Output

MONITORED TIMES

DATE	START	END	LENGTH (MINS)
01 07 75	09 26	10 00	34
	10 06	11 54	108
	13 20	13 53	33
	20 16	-	-
02 07 75	-	14 05	1069
	20 56	-	-
03 07 75	-	14 00	1024
	18 15	-	-
04 07 75	-	11 38	1043
	12 13	14 02	109
	20 35	-	-
06 07 75	-	15 46	2591
	16 16	-	-

....

31 07 75	-	10 05	732
	10 07	10 41	34
	10 45	10 54	9
	12 52	13 03	11

TOTAL TIME MONITORED = 31504 MINS

= 70.57%

USER NAMES

IDENT	USERNAMES				
PTK	KIRSSTEI	KIRSTEIN	IR	IRSTEIN	KIRS
	KIRS	KIRSTEI	GOULD	%OHI	
BRE	EEEEEEEE	MAHCN	MMAHON	MURPHY	
EDIN	PRATT	EISENSTA	EIENSTAD		
BLL	MCUNCEY	BARRCW	FCRTUNE	L	
	RA	BIGGIES	47OMOUNC	YEE MANN	
	THOMPSON	OMOUNCEY	FFRANKIN	R	MERRY
BLRD	ERCROGER	ROGERS	ERUNNING		
BUPA	ECLARD	POLLARC	POLLAD	POLLARS	
EDM	THCMAS	¬THOMAS	47		
IIRS	MAHCN				
LEED	LEVER	KAMEEN	LEVE		
LOUG	Y47dF	BAKERBAK	BAKER		
MANC	RANGELEY	NEVILLE	EE147		
NIMR	SMITH	LOGCN			
RPMS	HILLS	PRES142H			
CVL	THCMPSON	EEEE			
WESS	FCRSEY	CARMEL			
NEUR	CAMPBELL	LCGIN			
THAM	MACE	66	66CROWE	THAMES	1
[illegible]	CURRY				

LOGIN NUMBER

IDENT	NO. OF LOGINS	TOTAL CONNECT TIMES	AV. TIME PER LOGIN	HOSTS USED
PTK	34	458	13.47	5
BRE	10	96	9.60	2
EDIN	21	293	13.95	3
BLL	105	1151	10.96	10
BLRD	9	53	5.88	1
BUPA	15	392	26.13	1
EDM	23	714	31.04	2
IIRS	1	1	1.00	1
LEED	9	86	9.55	1
LOUG	5	29	5.80	2
MANC	20	241	12.05	2
NIMR	8	79	9.87	1
RPMS	35	346	9.88	3
CVL	2	8	4.00	1
WESS	48	1010	21.04	4
NEUR	13	39	3.00	5
THAM	34	455	13.38	8
UKAR	5	79	15.80	1

CONNECT TIME MATRIX (TIMES IN MINS)

IDENTS	SRI	MIT	ISI	UC/B	134	NLM	TOTAL
PTK			000425				000425
BLRD						000053	000053
BLL	000003		000130	000027		000808	000968
EDIN			000008		001049		001057
KING		000188	000023				000211
LOUG			000007			000022	000029
WNSM			000026	000004		000365	000395
THAM	000355		000003				000358
TOTALS	000358	000188	000622	000031	001049	001248	003496

A1.6 Sample PDP-9A Raw Data

In order to save space, blank lines and prompt lines have been removed from this Appendix.

```
 NEW SYSTEM - ARPA CONSOLE  20 31  ON  MON  13 DEC 76
13 DEC 76 20 31 RBF:CCRAL BACK
                IHD:INT RESET
                DR1:CCNTACT WITH 360  MADE AT 20 31
          2C 32 RBF:RBE ERROR " OR
          20 37 ICP:ICP HOS 170  PT 06 CH A
                TEL:CHL A  ZZF  LCG F8  U1
          20 41 ICP:ICP HCS  12  PT 00 CH B
          20 42 TEL:CHL B  XX   LCG 1R  U2
          20 45 TEL:CHL B LCG OUT & CLS
          20 47 TEL:CHL A LOG OUT & CLS
13 DEC 76 21 10 ICP:ICP HCS  70  PT 00 CH A
                TEL:CHL A CLS
          21 21 ICP:ICP HCS 176  PT 03 CH A
                TEL:CHL A  MJW
          21 23 TEL:CHL A CLS
          21 47 ICP:ICP HOS  76  PT 00 CH A
          21 48 TEL:CHL A  GA   LCG ZT  U1
          21 49 TEL:CHL A LOG OUT & CLS
          21 57 ICP:ICP HOS  55  PT 15 CH A
13 DEC 76 22 01 TEL:CHL A  HSM  LOG BJ  U1
          22 04 ICP:ICP HCS 176  PT 03 CH B
                TEL:CHL A LOG OUT & CLS
          22 29 ICP:ICP HCS  09  PT 03 CH B
                TEL:CHL B  RW   LCG BJ  U1
          22 34 TEL:CHL B LCG CUT & CLS
          22 47 ICP:ICP HOS  76  PT 00 CH B
          22 48 TEL:CHL B  GA   LCG ZT  U1
          22 53 ICP:ICP HOS 241  PT 30 CH C
          22 54 TEL:CHL B LCG OUT & CLS
                ICP:FTP HOS  76  PT 00 CH B
          22 55 FTP: USER ZT  GETTING FILE  INDRA.IN581
                DNC:$*JOB 518 CN RM3.RD2  -- XFTP         ARPA
13 DEC 76 23 09 TEL:CHL D CLS
                TEL:CHL B CLS
          23 21 TEL:CHL A CLS
14 DEC 76 00 19 ICP:ICP HOS  70  PT 00 CH A
                TEL:CHL A CLS
          00 56 DR1:CCNTACT WITH 360  LCST AT 00 56
14 DEC 76 01 17 ICP:ICP HCS  70  PT 00 CH A
                TEL:CHL A CLS
          01 41 DR1:CCNTACT WITH 360  MADE AT 01 41
  *ERROR*    P   CALL       1   75112  237446   14640  202620
```

A1.7 Sample_PDP-9A_Tidied_Data

```
13 DEC 76 20 31 ***: PDP-9  UP
13 DEC 76 20 31 RRF:CCRAL BACK
13 DEC 76 20 31 IHD:INT RESET
13 DEC 76 20 31 DR1:CONTACT WITH 360  MADE AT 20 31
13 DEC 76 20 32 RRF:RRE ERROR " OR
13 DEC 76 20 37 ICP:ICP HCS 170  PT 06 CH A
13 DEC 76 20 37 TEL:CHL A  ZZF  LOG F8  01
13 DEC 76 20 41 ICP:ICP HCS  12  PT 00 CH B
13 DEC 76 20 42 TEL:CHL B  XX   LOG 1R  02
13 DEC 76 20 45 TEL:CHL B LOG OUT & CLS
13 DEC 76 20 47 TEL:CHL A LOG OUT & CLS
13 DEC 76 21 10 ICP:ICP HCS  70  PT 00 CH A
13 DEC 76 21 10 TEL:CHL A CLS
13 DEC 76 21 21 ICP:ICP HCS 176  PT 03 CH A
13 DEC 76 21 21 TEL:CHL A  MJW
13 DEC 76 21 23 TEL:CHL A CLS
13 DEC 76 21 47 ICP:ICP HOS  76  PT 00 CH A
13 DEC 76 21 48 TEL:CHL A  GA   LOG ZT  01
13 DEC 76 21 49 TEL:CHL A LOG OUT & CLS
13 DEC 76 21 57 ICP:ICP HCS  55  PT 15 CH A
13 DEC 76 22 01 TEL:CHL A  HSM  LOG BJ  01
13 DEC 76 22 04 ICP:ICP HOS 176  PT 03 CH B
13 DEC 76 22 04 TEL:CHL A LOG OUT & CLS
13 DEC 76 22 29 ICP:ICP HCS  09  PT 03 CH B
13 DEC 76 22 29 TEL:CHL B  RW   LOG BJ  01
13 DEC 76 22 34 TEL:CHL B LOG OUT & CLS
13 DEC 76 22 47 ICP:ICP HCS  76  PT 00 CH B
13 DEC 76 22 48 TEL:CHL B  GA   LOG ZT  01
13 DEC 76 22 53 ICP:ICP HOS 241  PT 30 CH C
13 DEC 76 22 54 TEL:CHL B LOG OUT & CLS
13 DEC 76 22 54 ICP:FTP HCS  76  PT 00 CH B
13 DEC 76 22 55 FTP: USER ZT  GETTING FILE  INDRA.IN581
13 DEC 76 22 55 DNC:$*JOB 518 ON RM3.RD2  -- XFTP      ARPA
13 DEC 76 23 09 TEL:CHL D CLS
13 DEC 76 23 09 TEL:CHL B CLS
13 DEC 76 23 21 TEL:CHL A CLS
14 DEC 76 00 19 ICP:ICP HOS  70  PT 00 CH A
14 DEC 76 00.19 TEL:CHL A CLS
14 DEC 76 00 56 DR1:CONTACT WITH 360  LOST AT 00 56
14 DEC 76 01 17 ICP:ICP HCS  70  PT 00 CH A
14 DEC 76 01 17 TEL:CHL A CLS
14 DEC 76 01 41 DR1:CONTACT WITH 360  MADE AT 01 41
14 DEC 76 01 58 ***: PDP-9  CRASHED
```

A1.8 Sample PDP-9A Compressed Data

```
13 12 76 20 31 #01 01
13 12 76 20 31 #02 03
13 12 76 20 31 #02 00
13 12 76 20 31 #02 02
13 12 76 20 37 #21 170 06 01 00                00 10 00
13 12 76 20 37 #11 01 00 00 00
13 12 76 20 37 #21 170 06 02 006 ZZF   F8  00 10 00
13 12 76 20 37 #11 01 01 00 00
13 12 76 20 41 #21 012 00 01 01                00 04 00
13 12 76 20 41 #11 02 01 00 00
13 12 76 20 42 #21 012 00 02 018 XX    1R  00 03 00
13 12 76 20 42 #11 02 02 00 00
13 12 76 20 42 #11 01 01 00 00
13 12 76 20 42 #11 00 00 00 00
13 12 76 21 10 #21 070 00 01 00                00 00 00
13 12 76 21 10 #11 01 00 00 00
13 12 76 21 10 #11 00 00 00 00
13 12 76 21 21 #21 176 03 01 00                00 02 00
13 12 76 21 21 #11 01 00 00 00
13 12 76 21 21 #11 00 00 00 00
13 12 76 21 47 #21 076 00 01 00                00 02 00
13 12 76 21 47 #11 01 00 00 00
13 12 76 21 48 #21 076 00 02 006 GA    ZT  00 01 00
13 12 76 21 48 #11 01 01 00 00
13 12 76 21 48 #11 00 00 00 00
13 12 76 21 57 #21 055 15 01 00                00 07 00
13 12 76 21 57 #11 01 00 00 00
13 12 76 22 01 #21 055 15 02 006 HSM   BJ  00 03 00
13 12 76 22 04 #21 176 03 01 01                00 01 00
13 12 76 22 04 #11 02 01 00 00
13 12 76 22 04 #11 01 00 00 00
13 12 76 22 29 #21 009 03 01 01                00 05 00
13 12 76 22 29 #11 02 00 00 00
13 12 76 22 29 #21 009 03 02 006 RW    BJ  00 05 00
13 12 76 22 29 #11 02 01 00 00
13 12 76 22 31 #11 01 00 00 00
13 12 76 22 47 #21 076 00 01 01                00 07 00
13 12 76 22 47 #11 02 00 00 00
13 12 76 22 48 #21 076 00 02 006 GA    ZT  00 06 00
13 12 76 22 48 #11 02 01 00 00
13 12 76 22 53 #21 241 30 01 02                00 22 00
13 12 76 22 53 #11 03 01 00 00
13 12 76 22 53 #11 02 00 00 00
13 12 76 22 54 #22 076 00 01 01                00 15 00
13 12 76 22 54 #11 03 00 00 00
13 12 76 22 55 #42 076 00 02 01        ZT  00 14 00 INDRA.IN581
```

```
13 12 76 23 10 #11 01 00 00 00
13 12 76 23 17 #11 00 00 00 00
14 12 76 00 19 #21 070 00 01 00               00 00 00
14 12 76 00 19 #11 01 00 00 00
14 12 76 00 19 #11 00 00 00 00
14 12 76 00 56 #04 02
14 12 76 01 17 #21 070 00 01 00               00 00 00
14 12 76 01 17 #11 01 00 00 00
14 12 76 01 17 #11 00 00 00 00
14 12 76 01 41 #02 02
14 12 76 01 58 #08 01
```

ACRONYMS

ARPA	Advanced Research Projects Agency
BL	British Library
BLL	British Library, Lending Division
BLRD	British Library, Research and Development Dept.
bps	Bits per second
CADC	Computer Aided Design Centre (Cambridge)
HEP	High Energy Physics
IMP	Interface Message Processor
INDRA	InterNetwork Display and Remote Access
ISI	Information Sciences Institute
MEDLINE	MEDLARS On-Line
NAM	Network Access Machine
NBS	National Bureau of Standards
NLM	National Library of Medicine
NMM	Network Measurement Machine
PSTN	Public Switched Telephone Network
RL	Rutherford Laboratory
RSRE	Royal Signals and Radar Establishment
SIMP	Satellite Interface Message Processor
STEIN	Short-Term Experimental Information Network
TIP	Terminal Interface Message Processor
UCL	University College London
UCLA	University of California at Los Angeles

REFERENCES

1. Abrams, M.D. et al., "The Network Measurement System", IEE Trans. Comm., 1976.

2. Bolt, Beranek and Newman, Quarterly Technical Report No. 2 (1 April - 3 June 1975), Report No. 3106, July 1975.

3. Burren, J. et al., "Electric Users Manual", RHEL-72-003, Rutherford Laboratory, 1972.

4. Hogben, D. et al., "OMNITAB II User's Reference Manual", NBS Technical Note 552, 1971. and Numerical Analysis", NBS Handbook 101, 1966.

5. Holmes, P.L., "An Approach to the Development of Library and Information Networks with Special Reference to the United Kingdom", AGARD-CP-158, p6-1, 1974.

6. Kenney, S.B., "Report and Evaluation of the ARPA Network via the UCL IIP", INDRA Note 528, Department of Statistics and Computer Science, UCL, 1976.

7. Kirstein, P.T., "Distributed Computer Networks", Nature, 257, pp 549-554, 1975.

8. Kirstein, P.T., "University College London ARPANET Project - Quarterly Status Report No. 10 (1 April - 30 June 1976)", INDRA TR-29, Department of Statistics and Computer Science, UCL, September 1976.

9. Kirstein, P.T., "University College London ARPANET Project - Annual Report (1 January 1976 - 31 December 1976)", INDRA TR-34, Department of Statistics and Computer Science, UCL, 1977.

10. Kleinrock L. and Naylor, W., "On Measured Behaviour of the ARPA Network", Proc. AFIPS NCC, Vol 43, pp 767-780, 1974.

11. Lethbridge, H.I., "Analysis of Interactive Terminal Conversations with the MEDLINE Data Base", INDRA Note 573, Department of Statistics and Computer Science, UCL, 1976.

12. Opderbeck H. and Kleinrock, L., "The Influence of Control Procedures on the Performance of Packet-Switched Networks", Proc. Nat. Telecomm. Conf., San Diego, pp 810-817, 1974.

13. Ornstein, S.M. et al., "The Terminal IMP for the ARPA

Computer Network", Proc. AFIPS SJCC, Vol 40, pp 243-254, 1972.

14. Roberts, L.G. and Wessler, B.D., "The ARPA Network", Computer-Communication Networks, Prentice-Hall, pp 485-500, 1973.

15. Stokes, A.V., "A Users' Guide to Babbage", INDRA TR-13, Department of Statistics and Computer Science UCL, 1974.

16. Stokes, A.V. and Higginson, P.L., "The Problems of Connecting Hosts into ARPANET", Proc. Eur. Conf. Comm. Networks, pp 25-33, 1975.

17. Stokes, A.V., "Software Development in a Network Environment Using A High-Level Language", Proc. IFIP Working Conference on Software for Minicomputers, 1975.

18. Stokes, A.V. et al., "Monitoring and Access Control of the London Node of ARPANET", Proc. AFIPS NCC, Vol 45, pp 597-603, 1976.

19. Stokes, A.V., "The Babbage System Library", INDRA TR-35, Department of Statistics and Computer Science, UCL, 1976.

20. Stokes, A.V., "XFSTATS - A User Guide", INDRA Note 562, Department of Statistics and Computer Science, UCL, 1976.

21. Stokes, A.V., "PDP-9(B) Log Analysis, July 1975 to July 1976", INDRA Note 568, Department of Statistics and Computer Science, UCL, 1976.

22. Stokes, A.V., "LOGA - A Program to Reduce PDP-9A Logs to a Canonical Form", INDRA Note 587, Department of Statistics and Computer Science, UCL, 1977.

23. Stokes, A.V., "PDP-9(A) Log Analysis, Jan to Dec 1976", INDRA Note 622, Department of Statistics and Computer Science, UCL, 1977.

24. Stokes, A.V., "Access to an Information Retrieval System via a Distributed Computer Network", INDRA TR-40 Department of Statistics and Computer Science, UCL, 1977.

25. Stokes, A.V., "A System Programmer's Guide to Babbage", INDRA TR-39, Department of Statistics and Computer Science, UCL, 1977.

26. Treadwell, J.W., 'Measurement Program on the PDP-9', INDRA

Note 565, Department of Statistics and Computer Science, UCL, 1976.

27. Watkins, S.W. and Abrams, M.D., "Interpretation of Data in the Network Measurement System", NBS Technical Note 897, 1976.

Acknowledgements

We wish to acknowledge support by the British Library under grant SI/G/093. Also, such research would not be feasible without support of the many bodies for the ARPA project at University College, in particular, the Science Research Council (B/RG/5981 and 67022), the US Advanced Research Projects Agency (N00014-74-C-0240 and N00014-77-B-0005) and the UK Ministry of Defence (AT/2047/064).

We also wish to acknowledge the contributions of many people of the INDRA group at University College to this project, in particular, to David Bates, Roger Gould and Peter Mott who did much of the coding and debugging of the programs mentioned.